

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:52:49 ; Search time 206.831 Seconds
(without alignments)
542.461 Million cell updates/sec

Title: US-09-719-277C-9

Perfect score: 1026
Sequence: 1 AQLINLKEKPNVTPTAAHRT.....WLCFALSQLRPSKCAITVG 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02 : *
1: uniprot_sprot : *
2: uniprot_trembl : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|-------|---------------------|
| 1 | 425 | 41.4 | 122 2 | Q00687 hepatitis c |
| 2 | 419 | 40.8 | 95 2 | Q68874 hepatitis c |
| 3 | 386 | 37.6 | 102 2 | Q68358 hepatitis c |
| 4 | 377 | 36.7 | 100 2 | Q68360 hepatitis c |
| 5 | 368 | 35.9 | 102 2 | Q68361 hepatitis c |
| 6 | 352 | 34.3 | 102 2 | Q68359 hepatitis c |
| 7 | 334 | 32.6 | 102 2 | Q68356 hepatitis c |
| 8 | 318 | 31.0 | 102 2 | Q68365 hepatitis c |
| 9 | 316 | 30.8 | 102 2 | Q68366 hepatitis c |
| 10 | 305.5 | 29.8 | 134 2 | Q706D2 hepatitis c |
| 11 | 305.5 | 29.8 | 134 2 | CA64584 hepatitis c |
| 12 | 293 | 28.6 | 102 2 | Q68367 hepatitis c |
| 13 | 282 | 27.5 | 71 2 | Q91AW2 hepatitis c |
| 14 | 276 | 26.9 | 102 2 | Q68368 hepatitis c |
| 15 | 269 | 26.2 | 71 2 | Q91AW1 hepatitis c |
| 16 | 268 | 26.1 | 102 2 | Q68363 hepatitis c |
| 17 | 248 | 24.2 | 80 2 | Q81289 hepatitis c |
| 18 | 202 | 19.7 | 119 2 | Q86686 hepatitis c |
| 19 | 158 | 15.4 | 53 2 | Q91KM9 hepatitis c |
| 20 | 120 | 11.7 | 108 2 | P87760 hepatitis c |
| 21 | 120 | 11.7 | 108 2 | P87761 hepatitis c |
| 22 | 120 | 11.7 | 210 2 | Q6NYL1 homo sapien |
| 23 | 120 | 11.7 | 210 2 | AAH6547 |
| 24 | 119.5 | 11.6 | 609 2 | Q8FM53 |
| 25 | 117 | 11.4 | 463 2 | Q6IRJ7 |
| 26 | 117 | 11.4 | 463 2 | AAH70896 |
| 27 | 115.5 | 11.3 | 463 2 | Q8VIN2 |
| 28 | 114.5 | 11.2 | 468 2 | Q6ZRW4 |
| 29 | 114.5 | 11.2 | 468 2 | BAC87194 |
| 30 | 114 | 11.1 | 415 2 | Q8CCV9 |
| 31 | 114 | 11.1 | 463 1 | ANX7_MOUSE |

| | | | | | |
|----|-------|------|--------|----------|---------------------|
| 32 | 114 | 11.1 | 463 2 | Q8BP75 | Q8BP75 mus musculus |
| 33 | 113.5 | 11.1 | 903 2 | Q82HR3 | Q82HR3 streptomyces |
| 34 | 113 | 11.0 | 1732 2 | Q09451 | Q09451 bonnemaisson |
| 35 | 111 | 10.8 | 463 2 | Q922A2 | Q922A2 mus musculus |
| 36 | 109.5 | 10.7 | 297 2 | Q96GM4 | Q96GM4 homo sapien |
| 37 | 108 | 10.5 | 719 2 | Q73RV2 | Q73RV2 mycobacteri |
| 38 | 108 | 10.5 | 719 2 | AA06876 | AA06876 mycobacte |
| 39 | 104.5 | 10.2 | 310 2 | Q9NX73 | Q9NX73 homo sapien |
| 40 | 104.5 | 10.2 | 524 2 | Q71E72 | Q71E72 homo sapien |
| 41 | 104.5 | 10.2 | 524 2 | AAQ08897 | AAQ08897 homo sapi |
| 42 | 104.5 | 10.2 | 545 2 | Q8NDH8 | Q8NDH8 homo sapien |
| 43 | 104.5 | 10.2 | 623 2 | Q8WZA9 | Q8WZA9 homo sapien |
| 44 | 104 | 10.1 | 237 2 | Q6ZS94 | Q6ZS94 homo sapien |
| 45 | 104 | 10.1 | 237 2 | BAC87060 | BAC87060 homo sapi |

ALIGNMENTS

| RESULT 1 | ID | Q00687 | PRELIMINARY; | PRT; | 122 AA. |
|---|--|--|--------------|------|---------|
| AC | Q00687 | Q00687 | | | |
| DT | 01-NOV-1996 | (TREMBLrel. 01, Created) | | | |
| DT | 01-NOV-1996 | (TREMBLrel. 01, Last sequence update) | | | |
| DT | 01-DEC-2001 | (TREMBLrel. 19, Last annotation update) | | | |
| DE | Core (Fragment). | | | | |
| GN | Name=core; | | | | |
| OS | Hepatitis C virus. | | | | |
| OC | Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; | | | | |
| OX | Hepatitis. | | | | |
| NCBI_TaxId=1103; | | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=92268871; PubMed=1316939; | | | | |
| RA | Chan S., McOmish F., Holmes E., Dow B., Peutherer J., Follett E., | | | | |
| RA | Yap P., Simmonds P., | | | | |
| RT | "Analysis of a new hepatitis C virus type and its phylogenetic | | | | |
| RT | relationship to existing variants." | | | | |
| RL | J. Gen. Virol. 73:1131-1141(1992). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=94270990; PubMed=7545932; | | | | |
| RA | Hotta H., Doi H., Hayashi T., Furutani M., Soemarto W., Mizokami M., | | | | |
| RA | Ohba K., Homma M., | | | | |
| RT | "Analysis of the core and E1 envelope region sequences of a novel | | | | |
| RT | variant of hepatitis C virus obtained in Indonesia." | | | | |
| RL | Arch. Virol. 136:53-62(1994). | | | | |
| RN | [3] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Chan S.-W., | | | | |
| RL | Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases. | | | | |
| DR | EMBL, D10123; BAA01000.1; -. | | | | |
| FT | NON_TER | 1 | | | |
| FT | NON_TER | 122 | | | |
| SQ | SEQUENCE | 122 AA; 12632 MW; FBSC7A7F76AE29AC CRC64; | | | |
| Query Match | | | | | |
| Best local similarity 41.4%; Score 425; DB 2; Length 122; | | | | | |
| Matches 85; Conservative 70.2%; Pred. No. 2.3e-22; | | | | | |
| Matches 85; Conservative 29; Indels 0; Gaps 0; | | | | | |
| QY | 6 | LKEKPNVTPTAAHRTLSRVAVRSIAEFTCCAGAPDWVCALGLPSGRSLVEGASLSP | 65 | | |
| DB | 1 | LKEKPNVTPTAAHRTLSRVAVRSIAEFTCCAGAPDWVCALGLPSGRSLVEGASLSP | 60 | | |
| QY | 66 | RIAGPAGPGLSGTLGSPMAKRVAGGDSGCPALGLGAPMTPGGPAIVWSSIP | 125 | | |
| DB | 61 | KRVGAKAGPGLSGTLGSPMAKRVAGGDSGCPALGLGAPMTPGGPAIVWSSIP | 120 | | |
| QY | 126 | R 126 | | | |
| DB | 121 | R 121 | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:53:34 ; Search time 39 Seconds
(without alignments)
481.084 Million cell updates/sec

Title: US-09-719-277C-9

Perfect score: 1026
Sequence: 1 AQLNLKPKPVPTAAHRT.....MLFCPALSQLRPSKCATLWG 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 419 | 40.8 | 95 | 2 S44213 | core protein - hep |
| 2 | 113 | 11.0 | 1732 | 2 T43026 | probable DNA-direc |
| 3 | 111 | 10.8 | 463 | 2 S29170 | annexin VII - mous |
| 4 | 99.5 | 9.7 | 784 | 2 A26601 | elastin precursor |
| 5 | 97 | 9.5 | 152 | 2 T34649 | hypothetical prote |
| 6 | 97 | 9.5 | 313 | 2 T22828 | hypothetical prote |
| 7 | 96 | 9.4 | 277 | 2 F84336 | hypothetical prote |
| 8 | 95.5 | 9.3 | 419 | 2 G70602 | hypothetical prote |
| 9 | 95.5 | 9.3 | 466 | 2 A36674 | transcription fact |
| 10 | 95.5 | 9.3 | 1419 | 2 A41182 | collagen alpha 1(I |
| 11 | 95.5 | 9.3 | 1487 | 2 B41182 | collagen alpha 1(I |
| 12 | 95.5 | 9.3 | 1504 | 2 T49896 | glycine/proline-ri |
| 13 | 95 | 9.3 | 517 | 2 T10927 | 3C3, 18c protein - |
| 14 | 95 | 9.3 | 1492 | 2 A40333 | collagen alpha 1'(|
| 15 | 94 | 9.2 | 314 | 2 F70766 | hypothetical prote |
| 16 | 94 | 9.2 | 356 | 2 T28227 | hypothetical prote |
| 17 | 94 | 9.2 | 1113 | 2 T14260 | period protein Per |
| 18 | 94 | 9.2 | 1115 | 2 T13955 | period protein Per |
| 19 | 93.5 | 9.1 | 1366 | 2 T35985 | probable large pro |
| 20 | 93 | 9.1 | 290 | 2 T23416 | hypothetical prote |
| 21 | 93 | 9.1 | 538 | 2 T27156 | hypothetical prote |
| 22 | 93 | 9.1 | 1487 | 1 CGH06C | collagen alpha 1(I |
| 23 | 93 | 9.1 | 13288 | 2 T03099 | mucin, submaxillar |
| 24 | 92.5 | 9.0 | 523 | 2 T36677 | probable secretory |
| 25 | 92.5 | 9.0 | 527 | 2 T37055 | probable oxidoredu |
| 26 | 92.5 | 9.0 | 637 | 2 T35608 | polyketide hydroxy |
| 27 | 92.5 | 9.0 | 930 | 2 A25923 | progestrone recep |
| 28 | 92 | 9.0 | 316 | 2 S08169 | collagen col-12 pr |
| 29 | 92 | 9.0 | 316 | 2 S08170 | collagen col-13 pr |

| | | | | | |
|----|------|-----|------|----------|--------------------|
| 30 | 92 | 9.0 | 479 | 1 A31753 | transcription fact |
| 31 | 92 | 9.0 | 1418 | 2 T45467 | collagen alpha 1(I |
| 32 | 91.5 | 8.9 | 234 | 2 S25757 | Ig lambda chain - |
| 33 | 91 | 8.9 | 143 | 2 E72699 | hypothetical prote |
| 34 | 91 | 8.9 | 460 | 2 T33110 | hypothetical prote |
| 35 | 91 | 8.9 | 560 | 2 A60164 | platelet membrane |
| 36 | 91 | 8.9 | 589 | 2 T29299 | hypothetical prote |
| 37 | 91 | 8.9 | 1049 | 1 CGB07S | collagen alpha 1(I |
| 38 | 90.5 | 8.8 | 459 | 2 T35317 | probable serine/th |
| 39 | 90.5 | 8.8 | 1464 | 2 S59856 | collagen alpha 1(I |
| 40 | 90 | 8.8 | 770 | 1 S30293 | transcription fact |
| 41 | 90 | 8.8 | 775 | 1 EDBE11 | immediate-early tr |
| 42 | 90 | 8.8 | 892 | 2 T09071 | SH3 domains-co |
| 43 | 90 | 8.8 | 1042 | 1 CGGH1S | collagen alpha |
| 44 | 89.5 | 8.7 | 304 | 2 T22602 | hypothetical p |
| 45 | 89.5 | 8.7 | 411 | 2 S41945 | hypothetical F |

ALIGNMENTS

RESULT 1
S44213
core protein - hepatitis C virus
C/Species: hepatitis C virus
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-;
C/Accession: S44213
Kiferuch, H.H.
submitted to the EMBL Data Library, April 1994
A/Reference number: S44213
A/Accession: S44213
A/Status: preliminary
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-95 <PEU>
A/Accession: UNIPROT:069977 EMBL:X789950; NID:g475172; PIDN:CA55547.1; PID:g8605

Query Match 40.8%; Score 419; DB 2; Length 95;
Best Local Similarity 84.2%; Pred. No. 2.4e-25;
Matches 80; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

| | | |
|----|----|--|
| QY | 10 | PVNTPTAARITSSRAVRSLSAEFTCCAGADWYCARIGRPSGRSLVEGSLSPRING 69 |
| DB | 1 | PVNTPTAARITSSRAVRSLSAEFTCCAGADWYCARIGRPSGRSLVEGSLSPRING 69 |
| QY | 70 | PRAGPGLSPGTLGPMARVAGGQDGSCEPALGL 104 |
| DB | 61 | PRAGPGLSPGTLGPMARVAGGQDGSCEPALGL 95 |

RESULT 2
T43026
probable DNA-directed RNA polymerase (BC 2.7.7.6) II largest chain - Bonnemaisonia hamif
C/Species: Bonnemaisonia hamifera
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T43026
R/Stiller, J.W.; Hall, B.D.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z22292
A/Accession: T43026
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1732 <STI>
A/Cross-references: UNIPROT:009451; EMBL:U90209; NID:g3172133; PID:g3172134; PIDN:AAC184
C/Genetic: fph1
A/Genes: fph1
C/Superfamily: human DNA-directed RNA polymerase II largest chain
C/Keywords: DNA binding; nucleotidyltransferase

Query Match 11.0%; Score 113; DB 2; Length 1732;
Best Local Similarity 31.2%; Pred. No. 0.74;
Matches 54; Conservative 13; Mismatches 62; Indels 44; Gaps 10;

QY 48 IGRLPGRSLVGRASL-SP-----RIAGPRAGPL-SPGTLGPMARVAGGQD 94

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:56:09 ; Search time 152.494 Seconds
(without alignments)
414.586 Million cell updates/sec

Title: US-09-719-277C-9

Perfect score: 1026

Sequence: 1 AQLNLKKEPNVTPTAAHRT.....WLCFALQRLPSKCAITWG 195

Scoring table: BLOSUM62

Searched: 1370721 seqs, 324215800 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 1026 | 100.0 | 195 | 16 US-10-601-020-9 | Sequence 9, Appl1 |
| 2 | 774 | 75.4 | 195 | 16 US-10-601-020-13 | Sequence 13, Appl1 |
| 3 | 745.5 | 72.7 | 196 | 16 US-10-601-020-2 | Sequence 2, Appl1 |
| 4 | 745.5 | 72.7 | 196 | 16 US-10-601-020-10 | Sequence 10, Appl1 |
| 5 | 732.5 | 71.4 | 192 | 16 US-10-601-020-17 | Sequence 17, Appl1 |
| 6 | 654 | 63.7 | 195 | 16 US-10-601-020-16 | Sequence 16, Appl1 |
| 7 | 640.5 | 62.4 | 160 | 14 US-10-150-283-1 | Sequence 1, Appl1 |
| 8 | 621.5 | 60.6 | 161 | 9 US-09-736-959A-3 | Sequence 3, Appl1 |
| 9 | 618.5 | 60.3 | 161 | 9 US-09-736-959A-7 | Sequence 7, Appl1 |
| 10 | 614.5 | 59.9 | 197 | 16 US-10-601-020-12 | Sequence 12, Appl1 |
| 11 | 613.5 | 59.8 | 161 | 9 US-09-736-959A-1 | Sequence 1, Appl1 |
| 12 | 613.5 | 59.8 | 161 | 9 US-09-736-959A-32 | Sequence 32, Appl1 |

| | | | | | |
|----|-------|------|-------|-------------------------|------------------------------------|
| 14 | 600.5 | 58.5 | 194 | 16 US-10-601-020-15 | Sequence 15, Appl1 |
| 15 | 586 | 57.1 | 143 | 9 US-09-736-959A-5 | Sequence 5, Appl1 |
| 16 | 579.5 | 56.5 | 194 | 16 US-10-601-020-14 | Sequence 14, Appl1 |
| 17 | 579 | 56.4 | 143 | 9 US-09-736-959A-11 | Sequence 11, Appl1 |
| 18 | 574 | 55.9 | 143 | 9 US-09-736-959A-8 | Sequence 8, Appl1 |
| 19 | 573 | 55.8 | 143 | 9 US-09-736-959A-10 | Sequence 10, Appl1 |
| 20 | 567 | 55.3 | 143 | 9 US-09-736-959A-4 | Sequence 4, Appl1 |
| 21 | 561 | 54.7 | 143 | 9 US-09-736-959A-12 | Sequence 12, Appl1 |
| 22 | 561 | 54.7 | 192 | 16 US-10-601-020-11 | Sequence 9, Appl1 |
| 23 | 558 | 54.4 | 143 | 9 US-09-736-959A-9 | Sequence 11, Appl1 |
| 24 | 511.5 | 49.9 | 154 | 9 US-09-736-959A-16 | Sequence 16, Appl1 |
| 25 | 455 | 44.3 | 125 | 9 US-09-736-959A-14 | Sequence 14, Appl1 |
| 26 | 449.5 | 43.8 | 154 | 9 US-09-736-959A-15 | Sequence 15, Appl1 |
| 27 | 449 | 43.8 | 125 | 9 US-09-736-959A-20 | Sequence 20, Appl1 |
| 28 | 448 | 43.7 | 139 | 9 US-09-736-959A-17 | Sequence 17, Appl1 |
| 29 | 431 | 42.0 | 139 | 9 US-09-736-959A-6 | Sequence 6, Appl1 |
| 30 | 400 | 39.0 | 125 | 9 US-09-736-959A-21 | Sequence 21, Appl1 |
| 31 | 394 | 38.4 | 125 | 9 US-09-736-959A-19 | Sequence 19, Appl1 |
| 32 | 391 | 38.1 | 125 | 9 US-09-736-959A-13 | Sequence 13, Appl1 |
| 33 | 362 | 35.3 | 125 | 9 US-09-736-959A-18 | Sequence 18, Appl1 |
| 34 | 360 | 35.1 | 115 | 10 US-09-873-224-148 | Sequence 148, App |
| 35 | 118.5 | 11.5 | 291 | 14 US-10-094-749-2648 | Sequence 2648, Ap |
| 36 | 117.5 | 11.5 | 606 | 15 US-10-276-774-1383 | Sequence 1383, Ap |
| 37 | 117.5 | 11.5 | 19723 | 15 US-10-084-846A-5 | Sequence 5, Appl1 |
| 38 | 116.5 | 11.4 | 181 | 16 US-10-437-963-142272 | Sequence 142272, A |
| 39 | 113.5 | 11.1 | 903 | 14 US-10-156-761-11093 | Sequence 11093, A |
| 40 | 111 | 10.8 | 19695 | 15 US-10-084-846A-3 | Sequence 3, Appl1 |
| 41 | 110.5 | 10.8 | 19608 | 15 US-10-084-846A-8 | Sequence 8, Appl1 |
| 42 | 109.5 | 10.7 | 19725 | 15 US-10-084-846A-4 | Sequence 4, Appl1 |
| 43 | 109 | 10.6 | 224 | 16 US-10-437-963-183009 | Sequence 183009, Sequence 7, Appl1 |
| 44 | 107 | 10.4 | 19652 | 14 US-10-084-846A-7 | Sequence 1, Appl1 |
| 45 | 104.5 | 10.2 | 623 | 15 US-10-182-232-1 | Sequence 1, Appl1 |

ALIGNMENTS

RESULT 1
US-10-601-020-9
; Sequence 9, Application US/10601020
; Publication No. US20040156862A1
; GENERAL INFORMATION:
; APPLICANT: Branch, Andrea D.
; APPLICANT: Walowski, Jose L.
; APPLICANT: Stump, Dechard D.
; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
; FILE REFERENCE: RII-003CPUSCN
; CURRENT APPLICATION NUMBER: US/10/601,020
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/719277
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/088670
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: US 60/089138
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: PCT/US99/12929
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 9
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-601-020-9

Query Match 100.0%; Score 1026; DB 16; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.8e-75;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQLNLKKEPNVTPTAAHRTSSRVAVSLAEFTCCRAADPWVCARLGRPSGSLVEG 60
DB 1 AQLNLKKEPNVTPTAAHRTSSRVAVSLAEFTCCRAADPWVCARLGRPSGSLVEG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:53:49 ; Search time 52.1461 Seconds
(without alignments)
247.996 Million cell updates/sec

Title: US-09-719-277C-9
Perfect score: 1026
Sequence: 1 AQLILKEKENTPTAAHRT.....WLFCPALSQLRPSKATLVG 195

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--|
| 1 | 360 | 35.1 | 115 | 4 | US-09-878-281A-148 Sequence 148, App |
| 2 | 116.5 | 11.4 | 220 | 4 | US-09-252-991A-16739 Sequence 16739, A |
| 3 | 111.5 | 10.9 | 237 | 4 | US-09-252-991A-18764 Sequence 18764, A |
| 4 | 109 | 10.6 | 162 | 4 | US-09-252-991A-18207 Sequence 18207, A |
| 5 | 108 | 10.5 | 416 | 4 | US-09-252-991A-30923 Sequence 30923, A |
| 6 | 106.5 | 10.4 | 339 | 4 | US-09-252-991A-28087 Sequence 28087, A |
| 7 | 106 | 10.3 | 239 | 4 | US-09-252-991A-30037 Sequence 30037, A |
| 8 | 105 | 10.2 | 171 | 4 | US-09-252-991A-23216 Sequence 23216, A |
| 9 | 105 | 10.2 | 256 | 4 | US-09-252-991A-25670 Sequence 25670, A |
| 10 | 104.5 | 10.2 | 798 | 4 | US-09-252-991A-23774 Sequence 23774, A |
| 11 | 103.5 | 10.1 | 273 | 4 | US-09-252-991A-28324 Sequence 28324, A |
| 12 | 103 | 10.0 | 621 | 4 | US-09-252-991A-19125 Sequence 19125, A |
| 13 | 102.5 | 10.0 | 201 | 4 | US-09-252-991A-16881 Sequence 16881, A |
| 14 | 102.5 | 10.0 | 303 | 4 | US-09-252-991A-19799 Sequence 19799, A |
| 15 | 102.5 | 10.0 | 460 | 4 | US-09-252-991A-31082 Sequence 31082, A |
| 16 | 101 | 9.8 | 178 | 4 | US-09-252-991A-24449 Sequence 24449, A |
| 17 | 100 | 9.7 | 178 | 4 | US-09-252-991A-23713 Sequence 23713, A |
| 18 | 100 | 9.7 | 266 | 4 | US-09-252-991A-20968 Sequence 20968, A |
| 19 | 99 | 9.6 | 245 | 4 | US-09-252-991A-27050 Sequence 27050, A |
| 20 | 99 | 9.6 | 1027 | 4 | US-09-252-991A-26216 Sequence 26216, A |
| 21 | 98.5 | 9.6 | 212 | 4 | US-09-252-991A-31371 Sequence 31371, A |
| 22 | 98.5 | 9.6 | 534 | 4 | US-09-252-991A-20468 Sequence 20468, A |
| 23 | 98 | 9.6 | 615 | 4 | US-09-252-991A-23203 Sequence 23203, A |
| 24 | 97.5 | 9.5 | 745 | 2 | US-09-010-928B-28 Sequence 28, Appl |
| 25 | 97.5 | 9.5 | 788 | 4 | US-09-252-991A-28171 Sequence 28171, A |
| 26 | 97.5 | 9.5 | 870 | 2 | US-09-010-928B-2 Sequence 2, Appl |
| 27 | 97 | 9.5 | 148 | 4 | US-09-252-991A-27964 Sequence 27964, A |

| | | | | | |
|----|------|-----|-----|---|--|
| 28 | 97 | 9.5 | 159 | 4 | US-09-489-039A-7393 Sequence 7393, Ap |
| 29 | 97 | 9.5 | 192 | 4 | US-09-252-991A-22703 Sequence 22703, A |
| 30 | 96.5 | 9.4 | 203 | 4 | US-09-252-991A-17732 Sequence 17732, A |
| 31 | 96.5 | 9.4 | 244 | 4 | US-09-252-991A-18465 Sequence 18465, A |
| 32 | 96 | 9.4 | 251 | 4 | US-09-252-991A-28124 Sequence 28124, A |
| 33 | 95.5 | 9.3 | 240 | 4 | US-09-252-991A-21852 Sequence 21852, A |
| 34 | 95.5 | 9.3 | 415 | 4 | US-09-252-991A-19206 Sequence 19206, A |
| 35 | 95 | 9.3 | 546 | 4 | US-09-252-991A-18637 Sequence 18637, A |
| 36 | 95 | 9.3 | 558 | 4 | US-09-667-135-31 Sequence 31, Appl |
| 37 | 94.5 | 9.2 | 239 | 4 | US-09-252-991A-21250 Sequence 21250, A |
| 38 | 94.5 | 9.2 | 249 | 4 | US-09-252-991A-19213 Sequence 19213, A |
| 39 | 94.5 | 9.2 | 623 | 1 | US-08-653-740-7 Sequence 7, Appl |
| 40 | 94.5 | 9.2 | 623 | 2 | US-09-073-594-7 Sequence 7, Appl |
| 41 | 94.5 | 9.2 | 623 | 3 | US-09-275-925-7 Sequence 7, Appl |
| 42 | 94.5 | 9.2 | 707 | 4 | US-09-266-965-101 Sequence 101, Appl |
| 43 | 94 | 9.2 | 133 | 4 | US-09-252-991A-17771 Sequence 17771, A |
| 44 | 94 | 9.2 | 194 | 4 | US-09-252-991A-19292 Sequence 19292, A |
| 45 | 94 | 9.2 | 232 | 4 | US-09-252-991A-21812 Sequence 21812, A |

ALIGNMENTS

RESULT 1
US-09-878-281A-148
Sequence 148, Application US/09878281A
Patent No. 6762024
GENERAL INFORMATION:
APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, p
FILE REFERENCE: 35
CURRENT APPLICATION NUMBER: US/09/878,281A
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 148
LENGTH: 115
TYPE: PRT
ORGANISM: hepatitis C virus
US-09-878-281A-148

Query Match 35.1%; Score 360; DB 4; Length 115;
Best Local Similarity 67.6%; Pred. No. 2.4e-27;
Matches 73; Conservative 4; Mismatches 29; Indels 2; Gaps 1;

| | | | |
|----|----|---|-----|
| QY | 7 | KEKPNTPTAAHRTLSRVAVRSIAFTCCRAAPDMWCARLGRISGVGASISPR | 66 |
| DB | 10 | KTRRNTNP--GHTTSSQAARSLVEFTCYHAGAPSWCVQCARLPBGRNLAVGANPSFG | 67 |
| QY | 67 | IAGPRAGPLSPGTLGPPMAMRVAGGODGSCPPALGLGAPWTPGGG | 114 |
| DB | 68 | RAEPFAGPLSPGTLGPPMAMRVAGGODGSCPPALARRGAOMTPPAG | 115 |

RESULT 2
US-09-252-991A-16739
Sequence 16739, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16739
LENGTH: 220

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:44:54 ; Search time 217.787 Seconds
(without alignments)
321.196 Million cell updates/sec

Title: US-09-719-277C-9
Perfect score: 1026
Sequence: 1 AQLINKKRNVTPTAAHRT.....MLFCPLSQLPSKCATLVG 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 743.5 | 72.5 | 198 | 3 | AAV44554 Hepatitis |
| 2 | 736.5 | 71.8 | 198 | 3 | AAV44558 Hepatitis |
| 3 | 734.5 | 71.6 | 198 | 3 | AAV44545 Hepatitis |
| 4 | 646.5 | 63.0 | 198 | 3 | AAV44557 Hepatitis |
| 5 | 640.5 | 62.4 | 160 | 8 | ADG22790 Hepatitis |
| 6 | 621.5 | 60.6 | 161 | 6 | ABU61866 HCV core |
| 7 | 618.5 | 60.3 | 161 | 6 | ABU61870 HCV core |
| 8 | 616.5 | 60.1 | 161 | 6 | ABU61865 HCV core |
| 9 | 613.5 | 59.8 | 161 | 6 | ABU61864 HCV core |
| 10 | 604 | 58.9 | 198 | 3 | AAV44553 Hepatitis |
| 11 | 600 | 58.5 | 198 | 3 | AAV44555 Hepatitis |
| 12 | 586 | 57.1 | 143 | 6 | ABU61868 HCV core |
| 13 | 579 | 56.4 | 143 | 6 | ABU61874 HCV core |
| 14 | 574.5 | 56.0 | 198 | 3 | AAV44556 Hepatitis |
| 15 | 574 | 55.9 | 143 | 6 | ABU61871 HCV core |
| 16 | 573 | 55.8 | 143 | 6 | ABU61872 HCV core |
| 17 | 567 | 55.3 | 143 | 6 | ABU61867 HCV core |
| 18 | 561 | 54.7 | 143 | 6 | ABU61875 HCV core |
| 19 | 560.5 | 54.6 | 198 | 3 | AAV44552 Hepatitis |
| 20 | 558 | 54.4 | 198 | 3 | AAV44554 Hepatitis |
| 21 | 511.5 | 49.9 | 154 | 6 | ABU61879 HCV core |
| 22 | 455 | 44.3 | 125 | 6 | ABU61877 HCV core |
| 23 | 449.5 | 43.8 | 154 | 6 | ABU61878 HCV core |
| 24 | 449 | 43.8 | 125 | 6 | ABU61883 HCV core |
| 25 | 448 | 43.7 | 139 | 6 | ABU61880 HCV core |

| | | | | | | |
|----|-------|------|-------|---|----------|--------------------|
| 26 | 431 | 42.0 | 139 | 6 | ABU61869 | ABU61869 HCV core |
| 27 | 400 | 39.0 | 125 | 6 | ABU61884 | ABU61884 HCV core |
| 28 | 394 | 38.4 | 125 | 6 | ABU61882 | ABU61882 HCV core |
| 29 | 391 | 38.1 | 125 | 6 | ABU61876 | ABU61876 HCV core |
| 30 | 362 | 35.3 | 125 | 6 | ABU61881 | ABU61881 HCV core |
| 31 | 360 | 35.1 | 115 | 2 | AA63351 | AA63351 Hepatitis |
| 32 | 302 | 29.4 | 91 | 8 | ADK01937 | ADK01937 Hepatitis |
| 33 | 296.5 | 28.9 | 76 | 8 | ADK02062 | ADK02062 Hepatitis |
| 34 | 292.5 | 28.5 | 76 | 8 | ADK01350 | ADK01350 Hepatitis |
| 35 | 239.5 | 23.3 | 69 | 8 | ADK01829 | ADK01829 Hepatitis |
| 36 | 229 | 22.3 | 58 | 8 | ADK01459 | ADK01459 Hepatitis |
| 37 | 222.5 | 21.7 | 76 | 8 | ADK01709 | ADK01709 Hepatitis |
| 38 | 118.5 | 11.5 | 291 | 6 | ADA55080 | ADA55080 Human pro |
| 39 | 117.5 | 11.5 | 606 | 7 | AB811013 | AB811013 Human mem |
| 40 | 116.5 | 11.4 | 420 | 7 | AB067993 | AB067993 Pseudomon |
| 41 | 114 | 11.1 | 463 | 7 | ADDA7587 | ADDA7587 Rat Prote |
| 42 | 113.5 | 11.1 | 19938 | 6 | AB898398 | AB898398 Streptomy |
| 43 | 112 | 10.9 | 19938 | 6 | ABP76678 | ABP76678 Streptomy |
| 44 | 111.5 | 10.9 | 297 | 7 | ABO70018 | ABO70018 Pseudomon |
| 45 | 111.5 | 10.9 | 3640 | 4 | ABG23029 | ABG23029 Novel hum |

ALIGNMENTS

| | | |
|----------|--|------------------------------------|
| RESULT 1 | AAV44554 | standard; protein; 198 AA. |
| ID | AAV44554 | |
| AC | AAV44554 | |
| XX | | |
| DT | 04-APR-2000 | (first entry) |
| XX | | |
| DE | Hepatitis C virus protein encoded by DNA isolate D14853. | |
| XX | | |
| KM | Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine; | |
| KM | prevention; diagnosis; therapeutic target; anti-HCV therapy; | |
| XX | HCV infection; DNA isolate D14853. | |
| OS | Hepatitis C virus. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | Misc-difference 143 | /label= unknown |
| FT | Misc-difference 161 | /note= "Corresponds to stop codon" |
| FT | Misc-difference 184 | /label= unknown |
| FT | Misc-difference 184 | /note= "Corresponds to stop codon" |
| XX | | |
| PN | WO963941-A2. | |
| XX | | |
| PD | 16-DEC-1999. | |
| XX | | |
| PF | 09-JUN-1999. | 99WO-US012929. |
| XX | | |
| PR | 09-JUN-1998. | 98US-0088670P. |
| XX | 11-JUN-1998. | 98US-0089138P. |
| PA | (BRAN/) BRANCH A D. | |
| PA | (MALE/) MALEMSKI J L. | |
| PA | (STUM/) STUMP D D. | |
| XX | | |
| PI | Branch AD, Malewski JL, Stump DD; | |
| XX | | |
| DR | WPI; 2000-126431/11. | |
| XX | | |
| PT | Novel Hepatitis C virus peptides useful in vaccine compositions, for | |
| XX | diagnosing HCV infection and as therapeutic agents. | |
| XX | | |
| PS | Example 1; Page 39-41; 50pp; English. | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:52:49 ; Search time 14.8494 Seconds
(without alignments)
542.461 Million cell updates/sec

Title: US-09-719-277C-6
Perfect score: 69
Sequence: 1 AAHTSSRAVVR 14

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt_02:*
1: uniprot_sprot:*
2: uniprot_tr embl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|------------------|----------------------|
| 1 | 60 | 87.0 | 95 2 Q68874 | Q68874 hepatitis C |
| 2 | 60 | 87.0 | 102 2 Q68361 | Q68361 hepatitis C |
| 3 | 60 | 87.0 | 102 2 Q68365 | Q68365 hepatitis C |
| 4 | 60 | 87.0 | 102 2 Q68366 | Q68366 hepatitis C |
| 5 | 60 | 87.0 | 102 2 Q68367 | Q68367 hepatitis C |
| 6 | 43 | 62.3 | 686 1 CHEA_RHOSH | CHEA_RHOSH |
| 7 | 43 | 62.3 | 686 2 CAB87134 | CAB87134 rhodobacter |
| 8 | 42 | 60.9 | 102 2 Q68358 | Q68358 hepatitis C |
| 9 | 42 | 60.9 | 259 2 Q88U08 | Q88U08 mus musculus |
| 10 | 42 | 60.9 | 260 2 Q88094 | Q88094 oryza sativ |
| 11 | 41 | 59.4 | 57 2 Q684H8 | Q684H8 oryza sativ |
| 12 | 41 | 59.4 | 102 2 Q68356 | Q68356 hepatitis C |
| 13 | 41 | 59.4 | 122 2 Q00687 | Q00687 hepatitis C |
| 14 | 40 | 58.0 | 387 2 Q87KU0 | Q87KU0 vibrio para |
| 15 | 39 | 56.5 | 279 2 Q96910 | Q96910 homo sapien |
| 16 | 39 | 56.5 | 321 2 Q84048 | Q84048 oryza sativ |
| 17 | 39 | 56.5 | 431 2 Q935K0 | Q935K0 salmonella |
| 18 | 39 | 56.5 | 439 2 Q6MXN4 | Q6MXN4 serratia ma |
| 19 | 39 | 56.5 | 439 2 Q9L516 | Q9L516 salmonella |
| 20 | 39 | 56.5 | 439 2 CAES1689 | CAES1689 serratia |
| 21 | 39 | 56.5 | 482 1 K2C8_HUMAN | P05187 homo sapien |
| 22 | 39 | 56.5 | 483 2 Q6GM70 | Q6GM70 homo sapien |
| 23 | 39 | 56.5 | 505 2 Q6PAC7 | Q6PAC7 homo sapien |
| 24 | 39 | 56.5 | 505 2 AAH63513 | AAH63513 homo sapi |
| 25 | 39 | 56.5 | 586 2 Q851F9 | Q851F9 oryza sativ |
| 26 | 39 | 56.5 | 768 2 Q7XU33 | Q7XU33 oryza sativ |
| 27 | 38 | 55.1 | 199 2 Q8LJ08 | Q8LJ08 oryza sativ |
| 28 | 38 | 55.1 | 213 2 Q6Z0B3 | Q6Z0B3 oryza sativ |
| 29 | 38 | 55.1 | 213 2 BAD05499 | BAD05499 oryza sat |
| 30 | 38 | 55.1 | 213 2 BAD05671 | BAD05671 oryza sat |
| 31 | 38 | 55.1 | 233 2 Q6J342 | Q6J342 vaccinia vi |

| | | | | |
|----|----|------|-------------------|--------------------|
| 32 | 38 | 55.1 | 233 2 AAT10400 | AAT10400 vaccinia |
| 33 | 38 | 55.1 | 233 2 AAT10586 | AAT10586 vaccinia |
| 34 | 38 | 55.1 | 268 2 Q6REK2 | Q6REK2 rhodococcus |
| 35 | 38 | 55.1 | 268 2 AAR90186 | AAR90186 rhodococc |
| 36 | 38 | 55.1 | 343 2 Q68109 | Q68109 rhizobium 1 |
| 37 | 38 | 55.1 | 541 2 Q9ZP10 | Q9ZP10 zea mays (m |
| 38 | 38 | 55.1 | 560 2 O51417 | O51417 borrelia bu |
| 39 | 38 | 55.1 | 694 2 Q83J39 | Q83J39 shigella fl |
| 40 | 38 | 55.1 | 806 2 Q9FIM2 | Q9FIM2 arabidopsis |
| 41 | 38 | 55.1 | 927 2 Q7QPP1 | Q7QPP1 giardia lam |
| 42 | 38 | 55.1 | 970 2 Q7QW55 | Q7QW55 giardia lam |
| 43 | 38 | 55.1 | 1037 1 YHIV_ECOLI | YHIV_ECOLI |
| 44 | 38 | 55.1 | 1037 2 Q8FC18 | Q8FC18 escherichia |
| 45 | 38 | 55.1 | 1037 2 Q8X291 | Q8X291 escherichia |

ALIGNMENTS

| | | | | |
|---|---|--------------|----------|-------------------------|
| RESULT 1 | | | | |
| ID | Q68874 | PRELIMINARY; | PRT; | 95 AA. |
| AC | Q68874; | | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Created) | | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Last sequence update) | | | |
| DT | 01-JUN-2003 (TREMBLrel. 24, Last annotation update) | | | |
| DE | Core protein (Fragment). | | | |
| OS | Hepatitis C virus. | | | |
| OC | Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; | | | |
| OC | Hepatitis C virus. | | | |
| OX | NCBI_TaxID=11103; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=95270681; PubMed=7751366; | | | |
| RA | Feucht H.H., Zoellner B., Polyska S., Laufs R.; | | | |
| RT | "Study on reliability of commercially available hepatitis C virus | | | |
| RT | antibody tests." | | | |
| RL | J. Clin. Microbiol. 33:620-624(1995). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Feucht H.; | | | |
| RL | Submitted (APR-1994) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; X78950; CAA5547.1; -. | | | |
| DR | PIR; S44213; S44213. | | | |
| FT | NON_TER | 1 | | |
| FT | NON_TER | 95 | 95 | |
| SO | SEQUENCE | 95 AA; | 9702 MW; | 647C80587C6F892F CRC64; |
| Query Match | | | | |
| Best Local Similarity 100.0%; Score 60; DB 2; Length 95; | | | | |
| Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| QY | 1 AAHTSSRAVVR 13 | | | |
| DB | 7 AAHTSSRAVVR 19 | | | |
| RESULT 2 | | | | |
| ID | Q68361 | PRELIMINARY; | PRT; | 102 AA. |
| AC | Q68361; | | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Created) | | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Last sequence update) | | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) | | | |
| DE | Core protein (Fragment). | | | |
| OS | Hepatitis C virus. | | | |
| OC | Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; | | | |
| OC | Hepatitis C virus. | | | |
| OX | NCBI_TaxID=11103; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=BUK3; | | | |
| RX | MEDLINE=96030859; PubMed=7595353; | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:53:34 ; Search time 2.8 Seconds

(without alignments)
481.084 Million cell updates/sec

Title: US-09-719-277C-6

Perfect score: 69

Sequence: 1 AAHRTSSRAVVR 14

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------|
| 1 | 60 | 87.0 | 95 | 2 | S44213 core protein - hep |
| 2 | 43 | 62.3 | 686 | 2 | S70180 chemotaxis protein |
| 3 | 39 | 56.5 | 483 | 2 | A34720 keratin 8, type II |
| 4 | 38 | 55.1 | 326 | 2 | A40120 DNA polymerase III |
| 5 | 38 | 55.1 | 560 | 2 | D70157 probable membrane |
| 6 | 38 | 55.1 | 1037 | 2 | S47734 probable transport |
| 7 | 38 | 55.1 | 1037 | 2 | C68024 probable transport |
| 8 | 38 | 55.1 | 1037 | 2 | B91178 amexin - pepper |
| 9 | 37 | 53.6 | 314 | 2 | S66274 amexin - pepper |
| 10 | 37 | 53.6 | 537 | 2 | T20525 probable anti-term |
| 11 | 36 | 52.2 | 90 | 2 | E90776 hypothetical prote |
| 12 | 36 | 52.2 | 316 | 2 | T44749 hypothetical prote |
| 13 | 36 | 52.2 | 316 | 2 | E70664 hypothetical prote |
| 14 | 36 | 52.2 | 388 | 2 | H82120 succinyl-coa synth |
| 15 | 36 | 52.2 | 711 | 2 | T30107 hypothetical prote |
| 16 | 36 | 52.2 | 1286 | 2 | T18734 hypothetical prote |
| 17 | 36 | 52.2 | 1414 | 2 | T33236 hypothetical prote |
| 18 | 35 | 50.7 | 81 | 4 | S12496 hypothetical prote |
| 19 | 35 | 50.7 | 143 | 2 | T34647 probable integral |
| 20 | 35 | 50.7 | 240 | 2 | E72740 hypothetical prote |
| 21 | 35 | 50.7 | 317 | 2 | C86479 probable amexin p |
| 22 | 35 | 50.7 | 365 | 2 | T25296 hypothetical prote |
| 23 | 35 | 50.7 | 370 | 1 | TVMVT1 transforming prote |
| 24 | 35 | 50.7 | 370 | 1 | TVMVT1 transforming prote |
| 25 | 35 | 50.7 | 370 | 1 | TVMST1 transforming prote |
| 26 | 35 | 50.7 | 655 | 2 | T34705 hypothetical prote |
| 27 | 35 | 50.7 | 675 | 2 | T32299 hypothetical prote |
| 28 | 34.5 | 50.0 | 198 | 2 | A36646 ORF protein - equ |
| 29 | 34.5 | 50.0 | 541 | 2 | T34850 probable acid-CoA |

| | | | | | |
|----|----|------|-----|---|----------------------------|
| 30 | 34 | 49.3 | 100 | 2 | H70586 hypothetical prote |
| 31 | 34 | 49.3 | 101 | 2 | D42074 hypothetical prote |
| 32 | 34 | 49.3 | 111 | 2 | A25573 minicircle a prote |
| 33 | 34 | 49.3 | 150 | 2 | H87682 Maoc family protei |
| 34 | 34 | 49.3 | 156 | 2 | H82496 ribose ABC transpo |
| 35 | 34 | 49.3 | 220 | 2 | AE3221 conserved hypothet |
| 36 | 34 | 49.3 | 231 | 2 | S62402 major facilitator |
| 37 | 34 | 49.3 | 279 | 2 | G83626 transcription regu |
| 38 | 34 | 49.3 | 302 | 2 | F85068 N7 like-protein (I |
| 39 | 34 | 49.3 | 359 | 2 | F85068 MHC class I protei |
| 40 | 34 | 49.3 | 370 | 1 | D22930 DNA repair and gen |
| 41 | 34 | 49.3 | 370 | 1 | C40585 recf protein - Sta |
| 42 | 34 | 49.3 | 389 | 2 | T02158 hypothetical prote |
| 43 | 34 | 49.3 | 425 | 2 | S75024 hypothetical prote |
| 44 | 34 | 49.3 | 446 | 2 | B70776 probable glin2 - My |
| 45 | 34 | 49.3 | 448 | 2 | A87113 glutamine synthase |

ALIGNMENTS

RESULT 1
S44213
core protein - hepatitis C virus
C/Species: hepatitis C virus
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S44213
R/Faucht, H.H.
submitted to the EMBL Data Library, April 1994
A/Reference number: S44213
A/Accession: S44213
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-95 <FEU>
A/Cross-references: UNIPROT:Q68874; EMBL:X78950; NID:G475172; PIDN:CA55547.1; PID:G860.

Query Match 87.0%; Score 60; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.0007; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0;
QY 1 AAHRTSSRAVVR 13
DB 7 AAHRTSSRAVVR 19

RESULT 2
S70180
chemotaxis protein cheA - Rhodobacter sphaeroides
N/Alternate names: histidine autokinase
C/Species: Rhodobacter sphaeroides
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C/Accession: S70180; S49211
R/Ward, M.J.; Bell, A.W.; Hamblin, P.A.; Packer, H.L.; Armitage, J.P.
Mol. Microbiol. 17, 357-366, 1995
A/Title: Identification of a chemotaxis operon with two che Y genes in Rhodobacter sphaeroides
A/Reference number: S70178; MUID:96079285; PMID:7494484
A/Accession: S70180
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-686 <MAR>
A/Cross-references: UNIPROT:Q53135; EMBL:X80027; NID:G510668; PIDN:CA56530.1; PID:G510
A/Experimental source: strain WS8-N
C/Genetics:
A/Gene: cheA
C/Superfamily: chemotaxis protein cheA
C/Keywords: autophosphorylation; chemotaxis; phosphohistidine; phosphoprotein; sensory
F49/Binding site: phosphate (His) (covalent) (by autophosphorylation) #status,predicted

Query Match 62.3%; Score 43; DB 2; Length 686;
Best Local Similarity 81.8%; Pred. No. 5.9; Mismatches 9; Conservative 0; Indels 0; Gaps 0;
QY 3 HRTSSRAVVR 13

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:56:09 ; Search time 10.9483 Seconds
(without alignments)
414.586 Million cell updates/sec

Title: US-09-719-277C-6

Perfect score: 69

Sequence: 1 AAHRTSSRAVRC 14

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|-------|-------------------|
| 1 | 69 | 100.0 | 14 | US-10-601-020-6 |
| 2 | 60 | 87.0 | 125 | US-09-736-959A-20 |
| 3 | 60 | 87.0 | 143 | US-09-736-959A-4 |
| 4 | 60 | 87.0 | 143 | US-09-736-959A-8 |
| 5 | 60 | 87.0 | 143 | US-09-736-959A-9 |
| 6 | 60 | 87.0 | 143 | US-09-736-959A-10 |
| 7 | 60 | 87.0 | 161 | US-09-736-959A-7 |
| 8 | 60 | 87.0 | 192 | US-10-601-020-17 |
| 9 | 56 | 81.2 | 143 | US-09-736-959A-5 |
| 10 | 56 | 81.2 | 143 | US-09-736-959A-12 |
| 11 | 56 | 81.2 | 194 | US-10-601-020-15 |
| 12 | 52 | 75.6 | 139 | US-09-736-959A-6 |
| 13 | 48 | 69.6 | 143 | US-09-736-959A-11 |

| | | | | | | |
|----|----|------|-----|----|----------------------|--------------------|
| 14 | 48 | 69.6 | 160 | 14 | US-10-150-283-1 | Sequence 1, Appl1 |
| 15 | 48 | 69.6 | 161 | 9 | US-09-736-959A-1 | Sequence 1, Appl1 |
| 16 | 48 | 69.6 | 161 | 9 | US-09-736-959A-3 | Sequence 3, Appl1 |
| 17 | 48 | 69.6 | 161 | 9 | US-09-736-959A-32 | Sequence 32, Appl1 |
| 18 | 48 | 69.6 | 196 | 16 | US-10-601-020-2 | Sequence 2, Appl1 |
| 19 | 48 | 69.6 | 196 | 16 | US-10-601-020-10 | Sequence 10, Appl1 |
| 20 | 47 | 68.1 | 13 | 16 | US-10-601-020-4 | Sequence 4, Appl1 |
| 21 | 46 | 66.7 | 195 | 16 | US-10-601-020-9 | Sequence 9, Appl1 |
| 22 | 46 | 66.7 | 139 | 16 | US-10-601-020-12 | Sequence 12, Appl1 |
| 23 | 42 | 60.9 | 154 | 9 | US-09-736-959A-17 | Sequence 17, Appl1 |
| 24 | 42 | 60.9 | 154 | 9 | US-09-736-959A-15 | Sequence 15, Appl1 |
| 25 | 42 | 60.9 | 195 | 16 | US-10-601-020-13 | Sequence 13, Appl1 |
| 26 | 42 | 60.9 | 269 | 16 | US-10-437-963-106382 | Sequence 106382, A |
| 27 | 41 | 59.4 | 131 | 15 | US-10-425-114-66272 | Sequence 66272, A |
| 28 | 41 | 59.4 | 154 | 9 | US-09-736-959A-16 | Sequence 16, Appl1 |
| 29 | 41 | 59.4 | 192 | 16 | US-10-601-020-11 | Sequence 11, Appl1 |
| 30 | 41 | 59.4 | 152 | 16 | US-10-437-963-121427 | Sequence 121427, A |
| 31 | 40 | 58.0 | 161 | 9 | US-09-736-959A-2 | Sequence 2, Appl1 |
| 32 | 39 | 56.5 | 30 | 15 | US-10-264-049-3167 | Sequence 3167, Ap |
| 33 | 39 | 56.5 | 50 | 15 | US-10-424-599-235701 | Sequence 235701, A |
| 34 | 39 | 56.5 | 57 | 14 | US-10-106-698-7838 | Sequence 7838, Ap |
| 35 | 39 | 56.5 | 67 | 15 | US-10-264-049-3204 | Sequence 3204, Ap |
| 36 | 39 | 56.5 | 81 | 15 | US-10-264-049-3162 | Sequence 3162, Ap |
| 37 | 39 | 56.5 | 86 | 15 | US-10-264-049-3226 | Sequence 3226, Ap |
| 38 | 39 | 56.5 | 146 | 16 | US-10-437-963-186856 | Sequence 186856, A |
| 39 | 39 | 56.5 | 344 | 16 | US-10-437-963-158653 | Sequence 158653, A |
| 40 | 39 | 56.5 | 396 | 15 | US-10-425-114-65832 | Sequence 65832, A |
| 41 | 39 | 56.5 | 403 | 14 | US-10-100-294A-33 | Sequence 33, Appl1 |
| 42 | 39 | 56.5 | 403 | 14 | US-10-100-294A-42 | Sequence 42, Appl1 |
| 43 | 39 | 56.5 | 422 | 9 | US-09-779-307-18 | Sequence 18, Appl1 |
| 44 | 39 | 56.5 | 482 | 9 | US-09-779-307-17 | Sequence 17, Appl1 |
| 45 | 39 | 56.5 | 482 | 15 | US-10-080-334-187 | Sequence 187, App |

ALIGNMENTS

RESULT 1
US-10-601-020-6
Sequence 6, Application US/10601020
Publication No. US2004015662A1
GENERAL INFORMATION:
APPLICANT: Branch, Andrea D.
APPLICANT: Walewski, Jose L.
APPLICANT: Stump, Deborah D.
TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
FILE REFERENCE: RIT-003CPOSCN
CURRENT FILING DATE: 2003-06-20
PRIOR FILING DATE: 2001-04-13
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/088670
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: US 60/089138
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: PCT/US99/12929
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 6
LENGTH: 14
TYPE: PRT
ORGANISM: Hepatitis C virus
US-10-601-020-6
Query Match 100.0%; Score 69; DB 16; Length 14;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAHRTSSRAVRC 14
DB 1 AAHRTSSRAVRC 14

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: October 27, 2004, 09:53:49 ; Search time 3.74382 Seconds
(without alignments)
247.996 Million cell updates/sec

Title: US-09-719-277C-6
Perfect score: 69
Sequence: 1 AAHTSSRAVRC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/6ackfilas1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 39 | 56.5 | 11 | 2 | US-08-960-128-5 |
| 2 | 39 | 56.5 | 459 | 4 | US-09-252-991A-23184 |
| 3 | 39 | 56.5 | 482 | 4 | US-09-538-092-858 |
| 4 | 39 | 56.5 | 483 | 4 | US-09-919-497-79 |
| 5 | 39 | 56.5 | 1467 | 4 | US-09-252-991A-17657 |
| 6 | 38 | 55.1 | 115 | 4 | US-09-878-281A-148 |
| 7 | 38 | 55.1 | 220 | 4 | US-09-252-991A-20618 |
| 8 | 38 | 55.1 | 306 | 4 | US-09-489-039A-9110 |
| 9 | 37 | 53.6 | 36 | 4 | US-09-270-767-35463 |
| 10 | 37 | 53.6 | 36 | 4 | US-09-270-767-50680 |
| 11 | 37 | 53.6 | 304 | 4 | US-09-252-991A-29991 |
| 12 | 36 | 52.2 | 184 | 4 | US-09-325-932A-66 |
| 13 | 36 | 52.2 | 198 | 4 | US-09-252-991A-18773 |
| 14 | 36 | 52.2 | 471 | 4 | US-09-252-991A-19101 |
| 15 | 36 | 52.2 | 1007 | 4 | US-09-252-991A-18614 |
| 16 | 36 | 52.2 | 1090 | 4 | US-09-252-991A-23149 |
| 17 | 35 | 50.7 | 124 | 4 | US-09-248-796A-15423 |
| 18 | 35 | 50.7 | 152 | 4 | US-09-270-767-39708 |
| 19 | 35 | 50.7 | 152 | 4 | US-09-270-767-54925 |
| 20 | 35 | 50.7 | 259 | 4 | US-09-252-991A-24558 |
| 21 | 35 | 50.7 | 290 | 4 | US-09-252-991A-22229 |
| 22 | 35 | 50.7 | 370 | 4 | US-09-417-039-3 |
| 23 | 35 | 50.7 | 444 | 4 | US-09-252-991A-17767 |
| 24 | 35 | 50.7 | 445 | 4 | US-09-252-991A-28348 |
| 25 | 35 | 50.7 | 929 | 4 | US-09-252-991A-19435 |
| 26 | 35 | 50.7 | 6396 | 4 | US-09-410-551B-72 |
| 27 | 35 | 50.7 | 6396 | 4 | US-09-940-316B-72 |

| | | | | | | |
|----|------|------|-----|---|----------------------|-------------------|
| 28 | 34.5 | 50.0 | 501 | 4 | US-09-252-991A-27850 | Sequence 27850, A |
| 29 | 34 | 49.3 | 103 | 4 | US-09-248-796A-22005 | Sequence 22006, A |
| 30 | 34 | 49.3 | 106 | 4 | US-09-252-991A-23013 | Sequence 23013, A |
| 31 | 34 | 49.3 | 148 | 4 | US-09-513-999C-7402 | Sequence 7402, Ap |
| 32 | 34 | 49.3 | 154 | 4 | US-09-673-395A-356 | Sequence 356, App |
| 33 | 34 | 49.3 | 168 | 4 | US-09-252-991A-23672 | Sequence 23672, A |
| 34 | 34 | 49.3 | 247 | 4 | US-09-252-991A-23672 | Sequence 23672, A |
| 35 | 34 | 49.3 | 279 | 1 | US-08-988-197-9 | Sequence 9, Appl |
| 36 | 34 | 49.3 | 279 | 4 | US-08-988-197-9 | Sequence 9, Appl |
| 37 | 34 | 49.3 | 291 | 4 | US-09-252-991A-28472 | Sequence 28472, A |
| 38 | 34 | 49.3 | 302 | 4 | US-09-903-814A-14 | Sequence 14, Appl |
| 39 | 34 | 49.3 | 309 | 4 | US-09-489-039A-13249 | Sequence 13249, A |
| 40 | 34 | 49.3 | 372 | 3 | US-09-134-001C-5163 | Sequence 5163, Ap |
| 41 | 34 | 49.3 | 414 | 4 | US-09-252-991A-30034 | Sequence 30034, A |
| 42 | 34 | 49.3 | 414 | 4 | US-09-252-991A-17463 | Sequence 17463, A |
| 43 | 34 | 49.3 | 418 | 4 | US-09-252-991A-29836 | Sequence 29836, A |
| 44 | 34 | 49.3 | 419 | 4 | US-09-252-991A-29836 | Sequence 29836, A |
| 45 | 34 | 49.3 | 511 | 4 | US-09-252-991A-26078 | Sequence 26078, A |

ALIGNMENTS

RESULT 1
US-08-960-128-5
Sequence 5, Application US/08960128
Patent No. 5951985
GENERAL INFORMATION:
APPLICANT: Butler, Sandra M.
APPLICANT: Pomato, Nicholas
APPLICANT: Bos, Ebo
APPLICANT: Hanna, Micheal G.
APPLICANT: Haspel, Martin V.
APPLICANT: Hoover, Herbert C.
TITLE OF INVENTION: Tumor Associated Epitopes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Akzo No. 5951985el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,591
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
TELEFAX: (301) 977-0847
INFORMATION FOR SEQ. ID NO. 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULAR TYPE: peptide
HYPOTHEICAL: NO
US-08-960-128-5
Query Match 56.5%; Score 39; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.53;

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:44:54 ; Search time 15.636 Seconds
(without alignments)
321.196 Million cell updates/sec

Title: US-09-719-277C-6

Perfect score: 69

Sequence: 1 AAHRTSSRAVRC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneeqp1980a:*
- 2: geneeqp1990a:*
- 3: geneeqp2000a:*
- 4: geneeqp2001a:*
- 5: geneeqp2002a:*
- 6: geneeqp2003a:*
- 7: geneeqp2003b:*
- 8: geneeqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 69 | 100.0 | 14 | AAV44549 | AAV44549 Hepatitis |
| 2 | 60 | 87.0 | 125 | ABU61883 | ABU61883 HCV core |
| 3 | 60 | 87.0 | 143 | ABU61867 | ABU61867 HCV core |
| 4 | 60 | 87.0 | 143 | ABU61873 | ABU61873 HCV core |
| 5 | 60 | 87.0 | 143 | ABU61871 | ABU61871 HCV core |
| 6 | 60 | 87.0 | 143 | ABU61872 | ABU61872 HCV core |
| 7 | 60 | 87.0 | 161 | ABU61870 | ABU61870 HCV core |
| 8 | 60 | 87.0 | 198 | AAV44558 | AAV44558 Hepatitis |
| 9 | 56 | 81.2 | 143 | ABU61868 | ABU61868 HCV core |
| 10 | 56 | 81.2 | 143 | ABU61875 | ABU61875 HCV core |
| 11 | 56 | 81.2 | 198 | AAV44556 | AAV44556 Hepatitis |
| 12 | 52 | 75.4 | 139 | ABU61869 | ABU61869 HCV core |
| 13 | 48 | 69.6 | 143 | ABU61874 | ABU61874 HCV core |
| 14 | 48 | 69.6 | 160 | ADG22790 | ADG22790 Hepatitis |
| 15 | 48 | 69.6 | 161 | ABU61866 | ABU61866 HCV core |
| 16 | 48 | 69.6 | 161 | ABU61864 | ABU61864 HCV core |
| 17 | 48 | 69.6 | 198 | AAV44545 | AAV44545 Hepatitis |
| 18 | 47 | 68.1 | 13 | AAV44547 | AAV44547 Hepatitis |
| 19 | 46 | 66.7 | 198 | AAV44553 | AAV44553 Hepatitis |
| 20 | 42 | 60.9 | 139 | ABU61880 | ABU61880 HCV core |
| 21 | 42 | 60.9 | 154 | ABU61878 | ABU61878 HCV core |
| 22 | 42 | 60.9 | 198 | AAV44554 | AAV44554 Hepatitis |
| 23 | 41 | 59.4 | 154 | ABU61879 | ABU61879 HCV core |
| 24 | 41 | 59.4 | 198 | AAV44552 | AAV44552 Hepatitis |
| 25 | 41 | 59.4 | 207 | ABU68796 | ABU68796 Drosophila |

| | | | | | | |
|----|----|------|------|---|----------|--------------------|
| 26 | 40 | 58.0 | 161 | 6 | ABU61865 | ABU61865 HCV core |
| 27 | 39 | 56.5 | 11 | 2 | AAW08921 | AAW08921 Tumour as |
| 28 | 39 | 56.5 | 30 | 5 | ABP42035 | ABP42035 Human ova |
| 29 | 39 | 56.5 | 57 | 4 | AAg77064 | AAg77064 Human col |
| 30 | 39 | 56.5 | 67 | 5 | ABP42072 | ABP42072 Human ova |
| 31 | 39 | 56.5 | 81 | 5 | ABP42030 | ABP42030 Human ova |
| 32 | 39 | 56.5 | 86 | 5 | ABP42094 | ABP42094 Human ova |
| 33 | 39 | 56.5 | 108 | 4 | ABG08245 | ABG08245 Novel hum |
| 34 | 39 | 56.5 | 398 | 6 | AAE38086 | AAE38086 Human cyt |
| 35 | 39 | 56.5 | 403 | 6 | ABU03038 | ABU03038 Mose 11p1 |
| 36 | 39 | 56.5 | 403 | 6 | ABU03035 | ABU03035 Mose 11p1 |
| 37 | 39 | 56.5 | 459 | 7 | ABO74438 | ABO74438 Pseudomon |
| 38 | 39 | 56.5 | 482 | 6 | AAE38082 | AAE38082 Human cyt |
| 39 | 39 | 56.5 | 483 | 5 | AAU84289 | AAU84289 Human end |
| 40 | 39 | 56.5 | 483 | 6 | ABR47511 | ABR47511 Breast ca |
| 41 | 39 | 56.5 | 483 | 6 | ABU57621 | ABU57621 Different |
| 42 | 39 | 56.5 | 483 | 6 | ADL11040 | ADL11040 Human CDN |
| 43 | 39 | 56.5 | 500 | 4 | ABG09414 | ABG09414 Novel hum |
| 44 | 39 | 56.5 | 513 | 4 | ABG08133 | ABG08133 Novel hum |
| 45 | 39 | 56.5 | 1467 | 7 | ABO68911 | ABO68911 Pseudomon |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAV44549 | AAV44549 standard; peptide; 14 AA. |
| ID | AAV44549 |
| XX | XX |
| AC | AAV44549; |
| XX | XX |
| DT | 04-APR-2000 (first entry) |
| XX | XX |
| DE | Hepatitis C virus antigenic peptide ARF #2. |
| XX | XX |
| KW | Hepatitis C virus; HCV; ARF #2; alternate reading frame; hepatitis C; |
| KW | antigenic peptide; immunogenic; vaccine; HCV infection; antibody; |
| KW | prevention; diagnosis; therapeutic target; anti-HCV therapy. |
| XX | XX |
| OS | Hepatitis C virus. |
| XX | XX |
| PN | MO9963941-A2. |
| XX | XX |
| PD | 16-DEC-1999. |
| XX | XX |
| PF | 09-JUN-1999; 99WO-US012929. |
| XX | XX |
| PR | 09-JUN-1998; 98US-008670P. |
| XX | XX |
| PA | (BRAN/) BRANCH A D. |
| XX | XX |
| PA | (WALE/) WALEWSKI J L. |
| XX | XX |
| PI | (STUM/) STUMP D D. |
| XX | XX |
| PI | Branch AD, Malewski JL, Stump DD; |
| XX | XX |
| DR | WPI; 2000-126431/11. |
| XX | XX |
| PT | Novel Hepatitis C virus peptides useful in vaccine compositions, for |
| PT | diagnosing HCV infection and as therapeutic agents. |
| XX | XX |
| PS | Claim 11; Page 44; 50pp; English. |
| XX | XX |
| CC | The present sequence is a Hepatitis C virus (HCV) antigenic peptide ARF |
| CC | #2. The novel HCV peptide is encoded by a reading frame +1 or +2 relative |
| CC | to the standard HCV open reading frame hence not derived from the |
| CC | standard HCV polypeptide. The peptide elicits an immune response in |
| CC | patients infected with HCV and are produced during HCV infection. The |
| CC | present sequence is used as an immunogen to generate antibodies against |
| CC | HCV protein which are useful for diagnosing HCV infection. The peptide is |
| CC | also useful in vaccine compositions for preventing HCV infection, and as |
| CC | a target for anti-HCV therapy |
| XX | XX |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:52:49 ; Search time 14.8494 Seconds
(without alignments)
542.461 Million cell updates/sec

Title: US-09-719-277C-5
Perfect score: 76
Sequence: 1 NLKRPNTPTAC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|-------|------------------------------|
| 1 | 45 | 59.2 | 117 2 | 037388 common chim |
| 2 | 44 | 57.9 | 465 2 | 08MZR5 ctenocephal |
| 3 | 44 | 57.9 | 571 2 | 09LH17 arabidopsis |
| 4 | 44 | 57.9 | 571 2 | 08W4P0 arabidopsis |
| 5 | 43 | 56.6 | 80 2 | 081289 hepatitis c |
| 6 | 43 | 56.6 | 102 2 | 068358 hepatitis c |
| 7 | 43 | 56.6 | 240 2 | 09XRY5 nicotiana t |
| 8 | 43 | 56.6 | 275 2 | 06E7S1 oryza sativ |
| 9 | 43 | 56.6 | 375 2 | 0943R6 oryza sativ |
| 10 | 43 | 56.6 | 421 2 | 06HP43 bacillus th |
| 11 | 43 | 56.6 | 421 2 | 073EAL bacillus ce |
| 12 | 43 | 56.6 | 421 2 | 0811J4 bacillus ce |
| 13 | 43 | 56.6 | 421 2 | 0812B2 bacillus an |
| 14 | 43 | 56.6 | 421 2 | AAS39333 bacillus |
| 15 | 43 | 56.6 | 421 2 | AAR29451 bacillus |
| 16 | 43 | 56.6 | 585 2 | 06E822 oryza sativ |
| 17 | 43 | 56.6 | 680 2 | 08EAL7 sheanella |
| 18 | 43 | 56.6 | 743 1 | 08EAL7 sheanella |
| 19 | 43 | 56.6 | 790 2 | 0946L4 oryza sativ |
| 20 | 43 | 56.6 | 796 2 | 08MGR9 drosophila |
| 21 | 43 | 56.6 | 796 2 | 09VTR4 drosophila |
| 22 | 43 | 56.6 | 850 2 | 071103 bovine aden |
| 23 | 43 | 56.6 | 107 2 | 08BHZ2 mus musculu |
| 24 | 42 | 55.3 | 201 2 | 083P66 shigella fl |
| 25 | 42 | 55.3 | 231 2 | 07UMQ3 shigella fl |
| 26 | 42 | 55.3 | 260 2 | 070711 escherichia |
| 27 | 42 | 55.3 | 260 2 | CAE85174 escherich |
| 28 | 42 | 55.3 | 261 2 | 08PKR5 escherich |
| 29 | 42 | 55.3 | 275 2 | 09MG93 chrysodidym |
| 30 | 42 | 55.3 | 367 2 | 091696 xenopus lae |
| 31 | 42 | 55.3 | 492 1 | GLK1_TREMA 09x049 thermotoga |

| | | | | |
|----|----|------|--------|-------------------------------|
| 32 | 42 | 55.3 | 837 2 | 06TRH8 homo sapien |
| 33 | 42 | 55.3 | 837 2 | AAR26468 |
| 34 | 42 | 55.3 | 1113 1 | TDR7 RAT |
| 35 | 42 | 55.3 | 1955 1 | PC15 HUMAN |
| 36 | 41 | 53.9 | 208 2 | 07XIN5 |
| 37 | 41 | 53.9 | 414 2 | 07UBH8 |
| 38 | 41 | 53.9 | 632 2 | 07OWJ3 |
| 39 | 41 | 53.9 | 838 2 | 07EYL3 |
| 40 | 41 | 53.9 | 838 2 | BAC84340 |
| 41 | 41 | 53.9 | 1086 1 | TDR7 MOUSE |
| 42 | 40 | 52.6 | 135 2 | 091752 |
| 43 | 40 | 52.6 | 143 2 | 08NLC3 |
| 44 | 40 | 52.6 | 143 2 | CAP18959 |
| 45 | 40 | 52.6 | 157 1 | AZLB_BACSU 007920 bacillus su |

ALIGNMENTS

RESULT 1
ID 037388 PRELIMINARY; PRT; 117 AA.
AC 037388;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE E4.
OS Common chimpanzee papillomavirus type 1.
OC Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus; Common chimpanzee papillomavirus.
OX NCBI_Taxid=66261;
RN [1]
RP SEQUENCE FROM N.A.
RA Scitricariello F., Soza I., Braskey K.M., Hilliard J.K.;
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF020905; AAB71708.1; --
DR InterPro; IPR003861; Papilloma_E4.
DR Pfam; PF02711; Pap_E4; 1.
SQ SEQUENCE 117 AA; 13140 MW; 206FA34CB4F56158 CRC64;

Query Match 59.2%; Score 45; DB 2; Length 117;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 NLKRPNTPTAC 14
Db 70 NLKRPNTPTAC 82
(1)
RESULT 2
ID 08MZR5 PRELIMINARY; PRT; 465 AA.
AC 08MZR5;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Juvenile hormone epoxide hydrolase II.
OS Ctenocephalides felis (Cat flea)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Siphonaptera; Pulicidae; Pulicinae;
OC Ctenocephalides.
OX NCBI_Taxid=7515;
RN [1]
RP SEQUENCE FROM N.A.
RA Keiser K.C.L., Brandt K.S., Silver G.M., Wisniewski N.;
RT Cloning, Partial Purification and in vivo Developmental Profile of
RT Expression of the Juvenile Hormone Epoxide Hydrolase of
RT Ctenocephalides felis.
RL Arch. Insect Biochem. Physiol. 0:0-0(2002).
DR EMBL; AF503909; AAM22695.1; --
DR HSRP; Q9UR30; 1007.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR010497; EH_N.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using BW model

Run on: October 27, 2004, 09:53:34 ; Search time 2.8 seconds
(without alignments)
481.084 Million cell updates/sec

Title: US-09-719-277C-5
Perfect score: 76
Sequence: 1 LMLKEKPNVTPTAC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----------|-----------------------|
| 1 | 42 | 55.3 | 367 | 2 JC4151 | activin beta D cha |
| 2 | 42 | 55.3 | 493 | 2 C72314 | glycerol kinase - |
| 3 | 40 | 52.6 | 135 | 2 S31682 | inhibin beta-A cha |
| 4 | 40 | 52.6 | 157 | 2 F69592 | azlBCD operon tran |
| 5 | 40 | 52.6 | 304 | 1 PMRTNB | Na+/K+-exchanging |
| 6 | 40 | 52.6 | 409 | 2 D90268 | conserved hypochet |
| 7 | 39 | 51.3 | 233 | 2 H86593 | YOP translocation |
| 8 | 39 | 51.3 | 233 | 2 D72030 | type III secretion |
| 9 | 39 | 51.3 | 306 | 2 G84729 | hypothetical prote |
| 10 | 39 | 51.3 | 559 | 2 T09135 | hypothetical prote |
| 11 | 39 | 51.3 | 782 | 2 T32155 | hypothetical prote |
| 12 | 39 | 51.3 | 1246 | 2 T01358 | hypothetical prote |
| 13 | 39 | 51.3 | 1333 | 2 S38635 | blastoplia polyprot |
| 14 | 39 | 51.3 | 2658 | 2 A86216 | core protein - hep |
| 15 | 38 | 50.0 | 95 | 2 S44213 | ribosomal protein |
| 16 | 38 | 50.0 | 103 | 2 G70159 | thioesterdohin (impor |
| 17 | 38 | 50.0 | 108 | 2 B84999 | hypothetical prote |
| 18 | 38 | 50.0 | 177 | 2 T25533 | probable Deor-fam1 |
| 19 | 38 | 50.0 | 261 | 2 AD0963 | Partition protein |
| 20 | 38 | 50.0 | 283 | 2 AB1433 | Partition protein |
| 21 | 38 | 50.0 | 283 | 2 AC1797 | Partition protein |
| 22 | 38 | 50.0 | 304 | 1 S09601 | Na+/K+-exchanging |
| 23 | 38 | 50.0 | 304 | 2 B93974 | hypothetical prote |
| 24 | 38 | 50.0 | 333 | 2 JC2432 | L-lactate dehydrog |
| 25 | 38 | 50.0 | 395 | 2 AH2243 | hypothetical prote |
| 26 | 38 | 50.0 | 720 | 1 G64836 | probable membrane |
| 27 | 38 | 50.0 | 720 | 2 D90759 | hypothetical prote |
| 28 | 38 | 50.0 | 720 | 2 B85623 | hypothetical prote |
| 29 | 38 | 50.0 | 736 | 2 G01522 | acidic 82 kDa prot |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 38 | 50.0 | 827 | 1 C0BYD1 | RED1 protein - yea |
| 31 | 38 | 50.0 | 2888 | 2 I49477 | alpha-A-crystallin |
| 32 | 37 | 48.7 | 2 | S18989 | famA protein - Str |
| 33 | 37 | 48.7 | 275 | 2 T30573 | XylS/AraC transcri |
| 34 | 37 | 48.7 | 321 | 2 E89816 | ribose-phosphate p |
| 35 | 37 | 48.7 | 383 | 2 T19735 | hypothetical prote |
| 36 | 37 | 48.7 | 402 | 2 A45056 | hypothetical prote |
| 37 | 37 | 48.7 | 427 | 2 T23954 | hypothetical prote |
| 38 | 37 | 48.7 | 427 | 2 AC1627 | N-acetylmutamoyl-L |
| 39 | 37 | 48.7 | 496 | 2 T15691 | hypothetical prote |
| 40 | 37 | 48.7 | 497 | 2 T29791 | hypothetical prote |
| 41 | 37 | 48.7 | 601 | 2 E95863 | alcohol dehydrogen |
| 42 | 37 | 48.7 | 651 | 2 A96591 | NPX1-related prote |
| 43 | 37 | 48.7 | 710 | 1 XH0FK | formaldehyde trans |
| 44 | 37 | 48.7 | 754 | 1 BABOH | peptide-aspartate |
| 45 | 37 | 48.7 | 880 | 2 S60137 | beta-N-acetylhexos |

ALIGNMENTS

RESULT 1

JC4151
activin beta D chain precursor - African clawed frog

C/Species: Xenopus laevis (African clawed frog)
C/Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C/Accession: JC4151

R/Oda, S.; Nishimatsu, S.; Murakami, K.; Ueno, N.

Biochem. Biophys. Res. Commun. 210, 581-588, 1995
A/Title: Molecular cloning and functional analysis of a new activin beta subunit: a doxi

A/Reference number: JC4151; PMID:95275314; PMID:7755637

A/Accession: JC4151

A/Molecule type: mRNA
A/Residues: 1-367 <ODA>

A/Cross-references: UNIPROT:Q91696; DDBJ:D49543; NID:9961512; PIDD:BA08494.1; PTD:99611

A/Experimental source: embryo

C/Superfamily: Inhibin
C/Keywords: glycoprotein; mesoderm
F/1-253/Domain: signal sequence #status predicted <SIG>
F/254-367/Product: activin beta D chain #status predicted <MAT>
F/64,155,161,208,230/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 55.3%; Score 42; DB 2; Length 367;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

| QY | 1 LMLKEKPNVT 10 |
|----|------------------|
| DB | 57 LMLKEKPNVT 66 |

RESULT 2

C72314
glycerol kinase - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

R/Nelson, K.B.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.

Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A/Reference number: A72200; PMID:99287316; PMID:10360571

A/Accession: C72314

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-492 <ARN>

A/Cross-references: UNIPROT:Q9X049; GB:AE001758; GB:AE000512; NID:94981480; PIDD:AMD360;

A/Experimental source: strain MSB8

C/Genetics:

A/Gene: TM0952

C/Superfamily: xylulokinase

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:56:09 ; Search time 10.9483 Seconds
(without alignments)
414.586 Million cell updates/sec

Title: US-09-719-277C-5
Perfect score: 76
Sequence: 1 LNLKRPNTPTAC 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 76 | 100.0 | 14 | US-10-601-020-5 | Sequence 5, Appl1 |
| 2 | 67 | 88.2 | 192 | US-10-601-020-17 | Sequence 17, Appl1 |
| 3 | 67 | 88.2 | 195 | US-10-601-020-9 | Sequence 9, Appl1 |
| 4 | 67 | 88.2 | 195 | US-10-601-020-13 | Sequence 13, Appl1 |
| 5 | 67 | 88.2 | 195 | US-10-601-020-16 | Sequence 16, Appl1 |
| 6 | 63 | 82.9 | 196 | US-10-601-020-2 | Sequence 2, Appl1 |
| 7 | 63 | 82.9 | 196 | US-10-601-020-10 | Sequence 10, Appl1 |
| 8 | 63 | 82.9 | 196 | US-10-601-020-3 | Sequence 3, Appl1 |
| 9 | 51 | 67.1 | 160 | US-10-150-283-1 | Sequence 1, Appl1 |
| 10 | 51 | 67.1 | 194 | US-10-601-020-14 | Sequence 14, Appl1 |
| 11 | 47 | 61.8 | 192 | US-10-601-020-11 | Sequence 11, Appl1 |
| 12 | 47 | 61.8 | 197 | US-10-601-020-12 | Sequence 12, Appl1 |
| 13 | 45 | 59.2 | 993 | US-10-369-493-22570 | Sequence 22570, A |

| | | | | | | |
|----|----|------|------|----|----------------------|--------------------|
| 14 | 44 | 57.9 | 125 | 9 | US-09-736-959A-14 | Sequence 14, Appl1 |
| 15 | 44 | 57.9 | 143 | 9 | US-09-736-959A-4 | Sequence 4, Appl1 |
| 16 | 44 | 57.9 | 143 | 9 | US-09-736-959A-8 | Sequence 8, Appl1 |
| 17 | 44 | 57.9 | 143 | 9 | US-09-736-959A-9 | Sequence 9, Appl1 |
| 18 | 44 | 57.9 | 143 | 9 | US-09-736-959A-10 | Sequence 10, Appl1 |
| 19 | 44 | 57.9 | 143 | 9 | US-09-736-959A-11 | Sequence 11, Appl1 |
| 20 | 44 | 57.9 | 143 | 9 | US-09-736-959A-12 | Sequence 12, Appl1 |
| 21 | 44 | 57.9 | 161 | 9 | US-09-736-959A-7 | Sequence 7, Appl1 |
| 22 | 43 | 56.6 | 89 | 16 | US-10-437-963-161096 | Sequence 161096, A |
| 23 | 43 | 56.6 | 198 | 16 | US-10-767-701-35776 | Sequence 35776, A |
| 24 | 43 | 56.6 | 294 | 16 | US-10-437-963-204885 | Sequence 204885, A |
| 25 | 43 | 56.6 | 302 | 16 | US-10-437-963-138132 | Sequence 138132, A |
| 26 | 43 | 56.6 | 842 | 16 | US-10-437-963-138136 | Sequence 138136, A |
| 27 | 43 | 56.6 | 984 | 16 | US-10-437-963-204880 | Sequence 204880, A |
| 28 | 43 | 56.6 | 1102 | 16 | US-10-437-963-204882 | Sequence 204882, A |
| 29 | 42 | 55.3 | 96 | 16 | US-10-437-963-151268 | Sequence 151268, A |
| 30 | 42 | 55.3 | 260 | 14 | US-10-238-075-1141 | Sequence 1141, Ap |
| 31 | 42 | 55.3 | 492 | 14 | US-10-369-493-2984 | Sequence 2984, Ap |
| 32 | 42 | 55.3 | 1117 | 16 | US-10-250-615-12 | Sequence 12, Appl1 |
| 33 | 42 | 55.3 | 1955 | 14 | US-10-174-677-39 | Sequence 39, Appl1 |
| 34 | 42 | 55.3 | 1972 | 15 | US-10-085-198-20 | Sequence 20, Appl1 |
| 35 | 42 | 55.3 | 1973 | 15 | US-10-085-198-18 | Sequence 18, Appl1 |
| 36 | 41 | 53.9 | 65 | 15 | US-10-424-599-274401 | Sequence 274401, A |
| 37 | 41 | 53.9 | 150 | 15 | US-10-424-599-195434 | Sequence 195434, A |
| 38 | 41 | 53.9 | 378 | 16 | US-10-437-963-122759 | Sequence 122759, A |
| 39 | 41 | 53.9 | 498 | 16 | US-10-437-963-199013 | Sequence 199013, A |
| 40 | 41 | 53.9 | 606 | 16 | US-10-437-963-106462 | Sequence 106462, A |
| 41 | 40 | 52.6 | 53 | 11 | US-09-864-408A-1648 | Sequence 1648, Ap |
| 42 | 40 | 52.6 | 143 | 9 | US-09-738-626-6841 | Sequence 6841, Ap |
| 43 | 40 | 52.6 | 161 | 9 | US-09-738-959A-2 | Sequence 2, Appl1 |
| 44 | 40 | 52.6 | 161 | 9 | US-09-736-959A-3 | Sequence 3, Appl1 |
| 45 | 40 | 52.6 | 390 | 16 | US-10-437-963-110400 | Sequence 110400, A |

ALIGNMENTS

RESULT 1
US-10-601-020-5
Sequence 5, Application US/10601020
Publication No. US20040156862A1
GENERAL INFORMATION:
APPLICANT: Branch, Andrea D.
APPLICANT: Walewski, Jose L.
TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
FILE REFERENCE: RII-003CPUSCN
CURRENT APPLICATION NUMBER: US/10/601,020
CURRENT FILING DATE: 2003-06-20
PRIOR APPLICATION NUMBER: US 09/719277
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/088670
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: US 60/089138
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: PCT/US99/12929
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 14
TYPE: PRT
ORGANISM: Hepatitis C virus
US-10-601-020-5
Query Match 100.0%; Score 76; DB 16; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LNLKRPNTPTAC 14
DB 1 LNLKRPNTPTAC 14

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:53:49 ; Search time 3.74382 Seconds
(without alignments)
247.996 Million cell updates/sec

Title: US-09-719-277C-5

Perfect score: 76

Sequence: 1 LNLKRPVPTTAC 14

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCUTS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 44 | 57.9 | 430 | 3 | US-09-182-816-28 |
| 2 | 44 | 57.9 | 430 | 3 | US-09-471-528-28 |
| 3 | 44 | 57.9 | 430 | 3 | US-09-634-530-28 |
| 4 | 44 | 57.9 | 465 | 3 | US-09-182-816-23 |
| 5 | 44 | 57.9 | 465 | 3 | US-09-471-528-23 |
| 6 | 44 | 57.9 | 465 | 3 | US-09-634-530-23 |
| 7 | 42 | 55.3 | 145 | 4 | US-09-621-976-4002 |
| 8 | 40 | 52.6 | 342 | 4 | US-09-352-991A-23584 |
| 9 | 40 | 52.6 | 448 | 4 | US-09-520-781-32 |
| 10 | 40 | 52.6 | 590 | 4 | US-09-520-781-12 |
| 11 | 40 | 52.6 | 640 | 4 | US-09-907-794A-292 |
| 12 | 40 | 52.6 | 640 | 4 | US-09-905-125A-292 |
| 13 | 40 | 52.6 | 640 | 4 | US-09-902-775A-292 |
| 14 | 40 | 52.6 | 640 | 4 | US-09-906-700-292 |
| 15 | 40 | 52.6 | 640 | 4 | US-10-140-002-368 |
| 16 | 40 | 52.6 | 640 | 4 | US-09-303-603A-292 |
| 17 | 40 | 52.6 | 653 | 4 | US-09-520-781-10 |
| 18 | 40 | 52.6 | 653 | 4 | US-10-140-002-438 |
| 19 | 39 | 51.3 | 108 | 4 | US-09-198-452A-887 |
| 20 | 39 | 51.3 | 240 | 3 | US-09-329-418-7 |
| 21 | 39 | 51.3 | 240 | 3 | US-09-531-914-7 |
| 22 | 39 | 51.3 | 420 | 3 | US-09-329-418-8 |
| 23 | 39 | 51.3 | 420 | 3 | US-09-531-914-8 |
| 24 | 39 | 51.3 | 449 | 4 | US-09-270-767-5765 |
| 25 | 39 | 51.3 | 497 | 4 | US-09-345-473B-8 |
| 26 | 39 | 51.3 | 518 | 3 | US-09-329-418-3 |
| 27 | 39 | 51.3 | 518 | 3 | US-09-329-418-4 |

| | | | | | | |
|----|----|------|-----|---|----------------------|--------------------|
| 28 | 39 | 51.3 | 518 | 3 | US-09-329-418-5 | Sequence 5, Appl1 |
| 29 | 39 | 51.3 | 518 | 3 | US-09-329-418-9 | Sequence 9, Appl1 |
| 30 | 39 | 51.3 | 518 | 3 | US-09-531-914-3 | Sequence 3, Appl1 |
| 31 | 39 | 51.3 | 518 | 3 | US-09-531-914-4 | Sequence 4, Appl1 |
| 32 | 39 | 51.3 | 518 | 3 | US-09-531-914-5 | Sequence 5, Appl1 |
| 33 | 39 | 51.3 | 518 | 3 | US-09-531-914-9 | Sequence 9, Appl1 |
| 34 | 39 | 51.3 | 601 | 4 | US-09-270-767-42194 | Sequence 42194, A |
| 35 | 38 | 50.0 | 59 | 4 | US-09-461-325-481 | Sequence 481, App |
| 36 | 38 | 50.0 | 59 | 4 | US-10-012-542-481 | Sequence 481, App |
| 37 | 38 | 50.0 | 59 | 4 | US-10-115-123-481 | Sequence 481, App |
| 38 | 38 | 50.0 | 103 | 4 | US-09-732-210-741 | Sequence 741, App |
| 39 | 38 | 50.0 | 114 | 4 | US-09-461-325-479 | Sequence 479, App |
| 40 | 38 | 50.0 | 114 | 4 | US-10-012-542-479 | Sequence 479, App |
| 41 | 38 | 50.0 | 114 | 4 | US-10-115-123-479 | Sequence 479, App |
| 42 | 38 | 50.0 | 230 | 4 | US-09-252-991A-27570 | Sequence 27570, A |
| 43 | 38 | 50.0 | 367 | 4 | US-09-717-364A-25 | Sequence 25, Appl1 |
| 44 | 37 | 48.7 | 21 | 1 | US-07-956-700B-18 | Sequence 18, Appl1 |
| 45 | 37 | 48.7 | 21 | 1 | US-08-476-537-18 | Sequence 18, Appl1 |

ALIGNMENTS

RESULT 1
US-09-182-816-28
; Sequence 28, Application US/09182816
; Patent No. 6143542
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Io, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: NOVEL FLXA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
; FILE REFERENCE: FC-3-CI
; CURRENT APPLICATION NUMBER: US/09/182, 816
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989, 510
; EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-182-816-28
Query Match 57.9%; Score 44; DB 3; Length 430;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 LNLKRPVPTTAC 14
DB 353 LNLKRPVPTTAC 366
RESULT 2
US-09-471-528-28
; Sequence 28, Application US/09471528
; Patent No. 6153397
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Io, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: NOVEL FLXA EPOXIDE HYDROLASE PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-3-CI-1
; CURRENT APPLICATION NUMBER: US/09/471, 528
; CURRENT FILING DATE: 1999-12-27
; EARLIER APPLICATION NUMBER: 09/182, 816
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989, 510
; EARLIER FILING DATE: 1997-12-12

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:44:54 ; Search time 15.636 Seconds
(without alignments)
321.196 Million cell updates/sec

Title: US-09-719-277C-5

Perfect score: 76

Sequence: 1 LNLKPKNVTPTAC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1808:*
2: geneseqp1908:*
3: geneseqp2000:*
4: geneseqp2001:*
5: geneseqp2002:*
6: geneseqp2003:*
7: geneseqp2003b:*
8: geneseqp2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 76 | 100.0 | 14 | AAV44548 | Aay44548 Hepatitis |
| 2 | 67 | 88.2 | 198 | AAV44557 | Aay44557 Hepatitis |
| 3 | 67 | 88.2 | 198 | AAV44554 | Aay44554 Hepatitis |
| 4 | 67 | 88.2 | 198 | AAV44558 | Aay44558 Hepatitis |
| 5 | 63 | 82.9 | 198 | AAV44545 | Aay44545 Hepatitis |
| 6 | 51 | 67.1 | 13 | AAV44546 | Aay44546 Hepatitis |
| 7 | 51 | 67.1 | 160 | ADG22790 | Adg22790 Hepatitis |
| 8 | 51 | 67.1 | 198 | AAV44553 | Aay44553 Hepatitis |
| 9 | 47 | 61.8 | 198 | AAV44552 | Aay44552 Hepatitis |
| 10 | 47 | 61.8 | 125 | ABU61877 | Abu61877 HCV core |
| 11 | 44 | 57.9 | 143 | ABU61867 | Abu61867 HCV core |
| 12 | 44 | 57.9 | 143 | ABU61873 | Abu61873 HCV core |
| 13 | 44 | 57.9 | 143 | ABU61871 | Abu61871 HCV core |
| 14 | 44 | 57.9 | 143 | ABU61872 | Abu61872 HCV core |
| 15 | 44 | 57.9 | 143 | ABU61875 | Abu61875 HCV core |
| 16 | 44 | 57.9 | 143 | ABU61874 | Abu61874 HCV core |
| 17 | 44 | 57.9 | 161 | ABU61870 | Abu61870 HCV core |
| 18 | 44 | 57.9 | 430 | AAV24008 | Aay24008 A. flea ep |
| 19 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |
| 20 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |
| 21 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |
| 22 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |
| 23 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |
| 24 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |
| 25 | 44 | 57.9 | 430 | AAV24007 | Aay24007 A. flea ep |

ALIGNMENTS

| | | | | | | |
|----|----|------|------|---|----------|---------------------|
| 26 | 44 | 57.9 | 465 | 4 | AAB51009 | Aab51009 Flea epox |
| 27 | 44 | 57.9 | 571 | 8 | ADN72237 | Adn72237 Thale cre |
| 28 | 43 | 56.6 | 796 | 4 | ABB63128 | Abb63128 Drosophil |
| 29 | 43 | 56.6 | 1166 | 7 | ADF76887 | Adf76887 Novel hum |
| 30 | 43 | 56.6 | 1233 | 4 | ABB69686 | Abb69686 Drosophil |
| 31 | 42 | 55.3 | 145 | 6 | ABP76254 | Abp76254 Human GEN |
| 32 | 42 | 55.3 | 260 | 4 | ABB52866 | Abb52866 Escherich |
| 33 | 42 | 55.3 | 1117 | 5 | ABP53253 | Abp53253 Human MDD |
| 34 | 42 | 55.3 | 1972 | 5 | ADH48736 | Adh48736 NOVB8 pro |
| 35 | 42 | 55.3 | 1973 | 5 | ADH48734 | Adh48734 NOVB8 pro |
| 36 | 41 | 53.9 | 134 | 5 | ADK36761 | Adk36761 Novel hum |
| 37 | 40 | 52.6 | 53 | 5 | ADK31851 | Adk31851 Human ORF |
| 38 | 40 | 52.6 | 132 | 5 | ADK36831 | Adk36831 Novel hum |
| 39 | 40 | 52.6 | 143 | 4 | AAV23087 | Aay23087 C. glutami |
| 40 | 40 | 52.6 | 161 | 6 | ABU61866 | Abu61866 HCV core |
| 41 | 40 | 52.6 | 161 | 6 | ABU61865 | Abu61865 HCV core |
| 42 | 40 | 52.6 | 304 | 7 | ADK63893 | Adk63893 Rat Prote |
| 43 | 40 | 52.6 | 342 | 7 | ADK63889 | Adk63889 Rat Prote |
| 44 | 40 | 52.6 | 342 | 7 | ADK63889 | Adk63889 Rat Prote |
| 45 | 40 | 52.6 | 418 | 4 | AAU31980 | Aau31980 Pseudomon |

RESULT 1
ID AAY44548
AAV44548 standard; peptide; 14 AA.
AC AAY44548;
DT 04-APR-2000 (first entry)
DE Hepatitis C virus antigenic peptide ARF #1.
XX Hepatitis C virus; HCV; ARF #1; alternate reading frame; hepatitis C;
KW antigenic peptide; immunogenic; vaccine; HCV infection; antibody;
KW prevention; diagnosis; therapeutic target; anti-HCV therapy.
XX Hepatitis C virus.
OS Hepatitis C virus.
XX WO963941-A2.
XX 16-DEC-1999.
XX 09-JUN-1999; 99WO-US012929.
XX 09-JUN-1998; 98US-0086670P.
XX 11-JUN-1998; 98US-0089138P.
XX (BRAN/) BRANCH A D.
XX (WALE/) WALEWSKI J L.
XX (STUM/) STUMP D D.
XX Branch AD, Walewski JL, Stump DD;
XX WPI, 2000-126431/11.
XX Novel Hepatitis C virus peptides useful in vaccine compositions, for
XX diagnosing HCV infection and as therapeutic agents.
XX Claim 11; Page 44; 50pp; English.
XX The present sequence is a Hepatitis C virus (HCV) antigenic peptide ARF
XX #1. The novel HCV peptide is encoded by a reading frame +1 or +2 relative
XX to the standard HCV open reading frame hence not derived from the
XX standard HCV polypeptide. The peptide elicits an immune response in
XX patients infected with HCV and are produced during HCV infection. The
XX present sequence is used as an immunogen to generate antibodies against
XX HCV protein which are useful for diagnosing HCV infection. The peptide is
XX also useful in vaccine compositions for preventing HCV infection, and as
XX a target for anti-HCV therapy

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:52:49 ; Search time 13.7888 Seconds
(without alignments)
542.461 Million cell updates/sec

Title: US-09-719-277C-4
Perfect score: 51
Sequence: 1 AAHRTSSRXV 13

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: UniProt 02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 47 | 92.2 | 95 | 2 | Q68874 hepatitis c |
| 2 | 47 | 92.2 | 102 | 2 | Q68361 hepatitis c |
| 3 | 47 | 92.2 | 102 | 2 | Q68365 hepatitis c |
| 4 | 47 | 92.2 | 102 | 2 | Q68366 hepatitis c |
| 5 | 47 | 92.2 | 102 | 2 | Q68367 hepatitis c |
| 6 | 36 | 70.6 | 102 | 2 | Q68356 hepatitis c |
| 7 | 36 | 70.6 | 122 | 2 | Q00687 hepatitis c |
| 8 | 35 | 68.6 | 686 | 1 | Q68356 hepatitis c |
| 9 | 35 | 68.6 | 686 | 2 | Q68356 hepatitis c |
| 10 | 34 | 66.7 | 314 | 2 | Q68356 hepatitis c |
| 11 | 34 | 66.7 | 314 | 2 | Q68356 hepatitis c |
| 12 | 34 | 66.7 | 891 | 1 | Q68356 hepatitis c |
| 13 | 34 | 66.7 | 1319 | 2 | Q68356 hepatitis c |
| 14 | 34 | 66.7 | 1319 | 2 | Q68356 hepatitis c |
| 15 | 33 | 64.7 | 198 | 2 | Q68356 hepatitis c |
| 16 | 33 | 64.7 | 479 | 2 | Q68356 hepatitis c |
| 17 | 33 | 64.7 | 1467 | 2 | Q68356 hepatitis c |
| 18 | 33 | 64.7 | 1467 | 2 | Q68356 hepatitis c |
| 19 | 32 | 62.7 | 100 | 2 | Q68356 hepatitis c |
| 20 | 32 | 62.7 | 177 | 2 | Q68356 hepatitis c |
| 21 | 32 | 62.7 | 177 | 2 | Q68356 hepatitis c |
| 22 | 32 | 62.7 | 246 | 2 | Q68356 hepatitis c |
| 23 | 32 | 62.7 | 246 | 2 | Q68356 hepatitis c |
| 24 | 32 | 62.7 | 268 | 2 | Q68356 hepatitis c |
| 25 | 32 | 62.7 | 268 | 2 | Q68356 hepatitis c |
| 26 | 32 | 62.7 | 321 | 2 | Q68356 hepatitis c |
| 27 | 32 | 62.7 | 321 | 2 | Q68356 hepatitis c |
| 28 | 32 | 62.7 | 401 | 2 | Q68356 hepatitis c |
| 29 | 32 | 62.7 | 402 | 2 | Q68356 hepatitis c |
| 30 | 32 | 62.7 | 402 | 2 | Q68356 hepatitis c |
| 31 | 32 | 62.7 | 963 | 2 | Q68356 hepatitis c |

| | | | | | |
|----|----|------|-----|---|--------|
| 32 | 31 | 60.8 | 53 | 2 | Q68U90 |
| 33 | 31 | 60.8 | 129 | 2 | P71958 |
| 34 | 31 | 60.8 | 129 | 2 | Q71962 |
| 35 | 31 | 60.8 | 209 | 2 | Q81225 |
| 36 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 37 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 38 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 39 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 40 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 41 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 42 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 43 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 44 | 31 | 60.8 | 226 | 2 | Q7VU04 |
| 45 | 31 | 60.8 | 226 | 2 | Q7VU04 |

ALIGNMENTS

RESULT 1
ID Q68874 PRELIMINARY; PRT; 95 AA.
AC Q68874;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95270681; PubMed=7751366;
RA Feucht H.H., Zoellner B., Polymka S., Laufs R.;
RT "Study on reliability of commercially available hepatitis C virus
RT antibody tests";
RL J. Clin. Microbiol. 33:620-624 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Feucht H.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: X78950; CA55547.1; --
DR PIR: S44213; S44213.
FT NON_TER 1
FT NON_TER 95
SQ SEQUENCE 95 AA; 9702 MW; 647C80587C6F892F CRC64;
Query Match 92.2%; Score 47; DB 2; Length 95;
Best Local Similarity 76.9%; Pred. No. 0.0082;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAHRTSSRXV 13
Db 7 AAHRTSSRXV 19
RESULT 2
ID Q68361 PRELIMINARY; PRT; 102 AA.
AC Q68361;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EU03;
MEDLINE=96030659; PubMed=7595353;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:53:34 ; Search time 2.6 Seconds

(without alignments)
481.084 Million cell updates/sec

Title: US-09-719-277C-4

Perfect score: 51

Sequence: 1 AAHRTSSRXV 13

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------|
| 1 | 47 | 92.2 | 95 | 2 | S44213 core protein - hep |
| 2 | 35 | 68.6 | 686 | 2 | chemotaxis protein |
| 3 | 34 | 66.7 | 314 | 2 | annexin - pepper |
| 4 | 32 | 62.7 | 143 | 2 | glycoprotein D pre |
| 5 | 32 | 62.7 | 402 | 2 | probable glycoprote |
| 6 | 32 | 62.7 | 402 | 2 | hypothetical prote |
| 7 | 31 | 60.8 | 53 | 2 | hypothetical prote |
| 8 | 31 | 60.8 | 97 | 2 | hypothetical prote |
| 9 | 31 | 60.8 | 129 | 2 | hypothetical prote |
| 10 | 31 | 60.8 | 770 | 2 | hypothetical prote |
| 11 | 30 | 58.8 | 76 | 2 | protein gp31 - pha |
| 12 | 30 | 58.8 | 109 | 2 | hypothetical prote |
| 13 | 30 | 58.8 | 231 | 2 | major facilitator |
| 14 | 30 | 58.8 | 261 | 2 | hypothetical prote |
| 15 | 30 | 58.8 | 308 | 2 | annexin - alfalfa |
| 16 | 30 | 58.8 | 315 | 2 | annexin 2 - upland |
| 17 | 30 | 58.8 | 316 | 2 | annexin-like prote |
| 18 | 30 | 58.8 | 318 | 2 | annexin-like prote |
| 19 | 30 | 58.8 | 443 | 2 | hypothetical prote |
| 20 | 30 | 58.8 | 446 | 2 | probable gln2 - My |
| 21 | 30 | 58.8 | 448 | 2 | glutamine synthase |
| 22 | 30 | 58.8 | 463 | 2 | hypothetical prote |
| 23 | 30 | 58.8 | 561 | 2 | hypothetical prote |
| 24 | 30 | 58.8 | 571 | 2 | hypothetical prote |
| 25 | 30 | 58.8 | 614 | 2 | hypothetical prote |
| 26 | 30 | 58.8 | 743 | 2 | hypothetical prote |
| 27 | 30 | 58.8 | 748 | 2 | cholesterol O-acetyl |
| 28 | 30 | 58.8 | 916 | 2 | jockey protein 2 - |
| 29 | 29 | 56.9 | 103 | 2 | IS30 transposase I |

| | | | | | |
|----|----|------|-----|---|----------------------------|
| 30 | 29 | 56.9 | 103 | 2 | B85602 partial probable t |
| 31 | 29 | 56.9 | 117 | 2 | A98473 transposable membra |
| 32 | 29 | 56.9 | 251 | 2 | A91067 transposase of ins |
| 33 | 29 | 56.9 | 251 | 2 | IS30 transposase I |
| 34 | 29 | 56.9 | 279 | 2 | transcription regu |
| 35 | 29 | 56.9 | 303 | 2 | hypothetical prote |
| 36 | 29 | 56.9 | 315 | 2 | bifunctional cycla |
| 37 | 29 | 56.9 | 317 | 2 | probable annexin p |
| 38 | 29 | 56.9 | 343 | 2 | transposase of ins |
| 39 | 29 | 56.9 | 383 | 1 | transposase - Esch |
| 40 | 29 | 56.9 | 396 | 2 | hypothetical prote |
| 41 | 29 | 56.9 | 411 | 2 | transcription protein |
| 42 | 29 | 56.9 | 445 | 2 | hypothetical prote |
| 43 | 29 | 56.9 | 519 | 2 | transition protein |
| 44 | 29 | 56.9 | 549 | 2 | ig heavy chain pre |
| 45 | 29 | 56.9 | 663 | 2 | 78 kd glucose regu |

ALIGNMENTS

RESULT 1
S44213
core protein - hepatitis C virus
C/Species: hepatitis C virus
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S44213
R/Feucht, H.H.
submitted to the EMBL Data Library, April 1994
A/Reference number: S44213
A/Accession: S44213
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-95 <FEU>
A/Cross-references: UNIPROT:Q68874; EMBL:X78950; NID:G475172; PID:CAA5547.1; PID:98601

Query Match 92.2%; Score 47; DB 2; Length 95;
Best Local Similarity 76.9%; Pred. No. 0.003;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAHRTSSRXV 13
Db 7 AAHRTSSRXV 19

RESULT 2
S70180
chemotaxis protein cheA - Rhodospirillum rubrum
N/Alternate names: histidine autokinase
C/Species: Rhodospirillum rubrum
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C/Accession: S70180; S49211
R/Ward, M.J.; Bell, A.W.; Hamblin, P.A.; Packer, H.L.; Armitage, J.P.
Mol. Microbiol. 17, 357-366, 1995.
A/Title: Identification of a chemotaxis operon with two che Y genes in Rhodospirillum rubrum
A/Reference number: S70178; NID:96079255; PMID:749484
A/Accession: S70180
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-686 <WAR>
A/Cross-references: UNIPROT:Q53135; EMBL:X80027; NID:G510668; PID:CAA56330.1; PID:G5106
A/Experimental source: strain WS8-N
C/Genetics:
A/Gene: cheA
C/Superfamily: chemotaxis protein cheA
C/Keywords: autophosphorylation; chemotaxis; phosphohistidine; phosphoprotein; sensory t
F/49/Binding site: phosphate (His) (covalent) (by autophosphorylation) #status predicted

Query Match 68.6%; Score 35; DB 2; Length 686;
Best Local Similarity 63.6%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HRTSSRXV 13

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:56:09 ; Search time 10.1663 Seconds
(without alignments)
414.586 Million cell updates/sec

Title: US-09-719-277C-4

Perfect score: 51

Sequence: 1 AAHRTXSSRXVVR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/1/pubppa/US10D_PUBCOMB.pcp:*
17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pcp:*
19: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pcp:*
20: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 47 | 92.2 | 14 | 16 | US-10-601-020-6 |
| 2 | 47 | 92.2 | 125 | 9 | US-09-736-959A-20 |
| 3 | 47 | 92.2 | 143 | 9 | US-09-736-959A-4 |
| 4 | 47 | 92.2 | 143 | 9 | US-09-736-959A-8 |
| 5 | 47 | 92.2 | 143 | 9 | US-09-736-959A-9 |
| 6 | 47 | 92.2 | 143 | 9 | US-09-736-959A-10 |
| 7 | 47 | 92.2 | 161 | 9 | US-09-736-959A-7 |
| 8 | 47 | 92.2 | 192 | 16 | US-10-601-020-17 |
| 9 | 46 | 90.2 | 195 | 16 | US-10-601-020-9 |
| 10 | 46 | 90.2 | 197 | 16 | US-10-601-020-12 |
| 11 | 45 | 88.2 | 13 | 16 | US-10-601-020-4 |
| 12 | 44 | 86.3 | 139 | 9 | US-09-736-959A-6 |
| 13 | 43 | 84.3 | 143 | 9 | US-09-736-959A-5 |

| | | | | | | |
|----|----|------|-----|----|----------------------|--------------------|
| 14 | 43 | 84.3 | 143 | 9 | US-09-736-959A-12 | Sequence 12, Appl |
| 15 | 43 | 84.3 | 160 | 14 | US-10-150-283-1 | Sequence 1, Appl |
| 16 | 43 | 84.3 | 161 | 9 | US-09-736-959A-1 | Sequence 1, Appl |
| 17 | 43 | 84.3 | 161 | 9 | US-09-736-959A-3 | Sequence 3, Appl |
| 18 | 43 | 84.3 | 161 | 9 | US-09-736-959A-32 | Sequence 32, Appl |
| 19 | 43 | 84.3 | 194 | 16 | US-10-601-020-15 | Sequence 15, Appl |
| 20 | 43 | 84.3 | 196 | 16 | US-10-601-020-2 | Sequence 2, Appl |
| 21 | 43 | 84.3 | 196 | 16 | US-10-601-020-10 | Sequence 10, Appl |
| 22 | 42 | 82.4 | 195 | 16 | US-10-601-020-13 | Sequence 13, Appl |
| 23 | 40 | 78.4 | 143 | 9 | US-09-736-959A-11 | Sequence 11, Appl |
| 24 | 39 | 76.5 | 139 | 9 | US-09-736-959A-17 | Sequence 17, Appl |
| 25 | 39 | 76.5 | 154 | 9 | US-09-736-959A-15 | Sequence 15, Appl |
| 26 | 36 | 70.6 | 154 | 9 | US-09-736-959A-16 | Sequence 16, Appl |
| 27 | 36 | 70.6 | 192 | 16 | US-10-601-020-11 | Sequence 11, Appl |
| 28 | 36 | 70.6 | 788 | 14 | US-10-369-493-3863 | Sequence 3863, Ap |
| 29 | 35 | 68.6 | 115 | 10 | US-09-873-224-148 | Sequence 148, Ap |
| 30 | 35 | 68.6 | 161 | 9 | US-09-736-959A-2 | Sequence 2, Appl |
| 31 | 35 | 68.6 | 722 | 15 | US-10-320-727-3293 | Sequence 3293, Ap |
| 32 | 34 | 66.7 | 234 | 14 | US-10-029-368-33696 | Sequence 33696, A |
| 33 | 32 | 62.7 | 66 | 11 | US-09-864-408A-7018 | Sequence 7018, Ap |
| 34 | 32 | 62.7 | 124 | 16 | US-10-437-963-169292 | Sequence 169292, A |
| 35 | 32 | 62.7 | 329 | 14 | US-10-156-761-11490 | Sequence 11490, A |
| 36 | 32 | 62.7 | 346 | 16 | US-10-437-963-158653 | Sequence 158653, A |
| 37 | 31 | 60.8 | 52 | 16 | US-10-437-963-171059 | Sequence 171059, A |
| 38 | 31 | 60.8 | 59 | 15 | US-10-424-599-149827 | Sequence 149827, A |
| 39 | 31 | 60.8 | 78 | 15 | US-10-424-599-239876 | Sequence 239876, A |
| 40 | 31 | 60.8 | 90 | 9 | US-09-925-300-1168 | Sequence 1168, Ap |
| 41 | 31 | 60.8 | 96 | 16 | US-10-767-701-47836 | Sequence 47836, A |
| 42 | 31 | 60.8 | 102 | 15 | US-10-424-599-174229 | Sequence 174229, A |
| 43 | 31 | 60.8 | 110 | 16 | US-10-437-963-180838 | Sequence 180838, A |
| 44 | 31 | 60.8 | 226 | 15 | US-10-282-122A-51473 | Sequence 51473, A |
| 45 | 31 | 60.8 | 242 | 14 | US-10-393-840-55 | Sequence 55, Appl |

ALIGNMENTS

RESULT 1
US-10-601-020-6
Sequence 6, Application US/10601020
Publication No. US20040156862A1
GENERAL INFORMATION:
APPLICANT: Branch, Andrea D.
APPLICANT: Walewski, Jose L.
APPLICANT: Stump, Decker D.
TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
FILE REFERENCE: RIT-003CPUSCN
CURRENT APPLICATION NUMBER: US/10/601,020
CURRENT FILING DATE: 2003-06-20
PRIOR APPLICATION NUMBER: US 09/719277
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/088670
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: US 60/089138
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: PCT/US99/12929
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 14
TYPE: PRT
ORGANISM: Hepatitis C virus
US-10-601-020-6

Query Match 92.2% Score 47; DB 16; Length 14;
Best Local Similarity 76.9% Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAHRTXSSRXVVR 13
DB 1 AAHRTSSRAVVR 13

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:53:49 ; Search time 3.4764 seconds
(without alignments)
247.996 Million cell updates/sec

Title: US-09-719-277C-4
Perfect score: 51
Sequence: 1 AAHRTXSRRXVR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 10%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/1aa/5A COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B COMB.pep.*
3: /cgn2_6/prodata/1/1aa/6A COMB.pep.*
4: /cgn2_6/prodata/1/1aa/6B COMB.pep.*
5: /cgn2_6/prodata/1/1aa/PCUS COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 35 | 68.6 | 115 | 4 | US-09-878-281A-148 |
| 2 | 35 | 68.6 | 220 | 4 | US-09-252-991A-20618 |
| 3 | 35 | 68.6 | 1467 | 4 | US-09-252-991A-17657 |
| 4 | 33 | 64.7 | 418 | 4 | US-09-252-991A-17463 |
| 5 | 32 | 62.7 | 444 | 4 | US-09-252-991A-17767 |
| 6 | 32 | 62.7 | 894 | 4 | US-09-252-991A-30392 |
| 7 | 31 | 60.8 | 309 | 4 | US-09-489-039A-13237 |
| 8 | 31 | 60.8 | 401 | 4 | US-09-489-039A-13237 |
| 9 | 31 | 60.8 | 474 | 4 | US-09-252-991A-27524 |
| 10 | 30 | 58.8 | 116 | 4 | US-09-790-045-9 |
| 11 | 30 | 58.8 | 116 | 4 | US-10-222-577-9 |
| 12 | 30 | 58.8 | 116 | 4 | US-10-222-578-9 |
| 13 | 30 | 58.8 | 130 | 4 | US-09-489-039A-11402 |
| 14 | 30 | 58.8 | 175 | 4 | US-09-252-991A-33721 |
| 15 | 30 | 58.8 | 184 | 4 | US-09-325-932A-66 |
| 16 | 30 | 58.8 | 207 | 4 | US-09-489-039A-12450 |
| 17 | 30 | 58.8 | 295 | 4 | US-09-252-991A-21169 |
| 18 | 30 | 58.8 | 594 | 4 | US-10-140-002-10 |
| 19 | 30 | 58.8 | 630 | 4 | US-09-657-252-4 |
| 20 | 30 | 58.8 | 929 | 4 | US-09-252-991A-19435 |
| 21 | 29 | 56.9 | 76 | 4 | US-09-270-767-59232 |
| 22 | 29 | 56.9 | 107 | 4 | US-09-252-991A-27780 |
| 23 | 29 | 56.9 | 154 | 4 | US-09-252-991A-24472 |
| 24 | 29 | 56.9 | 191 | 4 | US-09-252-991A-28933 |
| 25 | 29 | 56.9 | 213 | 4 | US-09-252-991A-28496 |
| 26 | 29 | 56.9 | 284 | 4 | US-09-252-991A-25993 |
| 27 | 29 | 56.9 | 291 | 4 | US-09-252-991A-28472 |

| | | | | | | |
|----|----|------|------|---|----------------------|-------------------|
| 28 | 29 | 56.9 | 302 | 4 | US-09-903-814A-14 | Sequence 14, Appl |
| 29 | 29 | 56.9 | 349 | 4 | US-09-489-039A-8584 | Sequence 8584, Ap |
| 30 | 29 | 56.9 | 452 | 1 | US-08-336-618-23 | Sequence 23, Appl |
| 31 | 29 | 56.9 | 465 | 4 | US-09-252-991A-24848 | Sequence 24848, A |
| 32 | 29 | 56.9 | 522 | 4 | US-09-489-039A-10999 | Sequence 10999, A |
| 33 | 29 | 56.9 | 560 | 1 | US-08-336-618-22 | Sequence 22, Appl |
| 34 | 29 | 56.9 | 611 | 4 | US-09-252-991A-23379 | Sequence 23379, A |
| 35 | 29 | 56.9 | 711 | 4 | US-09-252-991A-22259 | Sequence 22259, A |
| 36 | 29 | 56.9 | 6396 | 4 | US-09-410-551B-72 | Sequence 72, Appl |
| 37 | 29 | 56.9 | 6396 | 4 | US-09-940-316B-72 | Sequence 72, Appl |
| 38 | 29 | 56.9 | 139 | 4 | US-09-252-991A-26906 | Sequence 26906, A |
| 39 | 28 | 54.9 | 146 | 3 | US-09-335-409-17 | Sequence 17, Appl |
| 40 | 28 | 54.9 | 146 | 3 | US-09-568-102-17 | Sequence 17, Appl |
| 41 | 28 | 54.9 | 146 | 3 | US-09-567-969-17 | Sequence 17, Appl |
| 42 | 28 | 54.9 | 146 | 3 | US-09-568-480-17 | Sequence 17, Appl |
| 43 | 28 | 54.9 | 146 | 3 | US-09-568-486-17 | Sequence 17, Appl |
| 44 | 28 | 54.9 | 146 | 3 | US-09-568-472-17 | Sequence 17, Appl |
| 45 | 28 | 54.9 | 146 | 3 | US-09-567-899-17 | Sequence 17, Appl |

ALIGNMENTS

RESULT 1
US-09-878-281A-148
; Sequence 148, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Immunogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 115
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-148

Query Match 68.6%; Score 35; DB 4; Length 115;
Best Local Similarity 63.6%; Pred. No. 1.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRTXSRRXVR 13
DB 19 HRTLSQAAVR 29

RESULT 2
US-09-252-991A-20618
; Sequence 20618, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20618
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20618

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:53:49 ; Search time 3.4764 Seconds
(without alignments) 247.996 Million cell updates/sec

Title: US-09-719-277C-4
Perfect score: 51
Sequence: 1 AAHRTXSRXXVR 13

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1 | 35 | 68.6 | 115 | 4 US-09-878-281A-148 | Sequence 148, App |
| 2 | 35 | 68.6 | 220 | 4 US-09-252-991A-20618 | Sequence 20618, A |
| 3 | 35 | 68.6 | 1467 | 4 US-09-252-991A-17657 | Sequence 17657, A |
| 4 | 33 | 64.7 | 448 | 4 US-09-252-991A-17463 | Sequence 17463, A |
| 5 | 32 | 62.7 | 894 | 4 US-09-252-991A-17677 | Sequence 17677, A |
| 6 | 32 | 62.7 | 894 | 4 US-09-252-991A-13097 | Sequence 13097, A |
| 7 | 31 | 60.8 | 309 | 4 US-09-252-991A-33092 | Sequence 33092, A |
| 8 | 31 | 60.8 | 401 | 4 US-09-489-039A-13337 | Sequence 13337, A |
| 9 | 31 | 60.8 | 474 | 4 US-09-252-991A-27524 | Sequence 27524, A |
| 10 | 30 | 58.8 | 116 | 4 US-09-790-045-9 | Sequence 9, Appl |
| 11 | 30 | 58.8 | 116 | 4 US-10-222-577-9 | Sequence 9, Appl |
| 12 | 30 | 58.8 | 116 | 4 US-10-222-577-9 | Sequence 9, Appl |
| 13 | 30 | 58.8 | 130 | 4 US-09-489-039A-11402 | Sequence 11402, A |
| 14 | 30 | 58.8 | 175 | 4 US-09-252-991A-23721 | Sequence 23721, A |
| 15 | 30 | 58.8 | 184 | 4 US-09-325-932A-66 | Sequence 66, Appl |
| 16 | 30 | 58.8 | 207 | 4 US-09-489-039A-1450 | Sequence 1450, A |
| 17 | 30 | 58.8 | 295 | 4 US-09-252-991A-21169 | Sequence 21169, A |
| 18 | 30 | 58.8 | 594 | 4 US-10-140-002-10 | Sequence 10, Appl |
| 19 | 30 | 58.8 | 630 | 4 US-09-657-252-4 | Sequence 4, Appl |
| 20 | 29 | 56.9 | 76 | 4 US-09-252-991A-19435 | Sequence 19435, A |
| 21 | 29 | 56.9 | 107 | 4 US-09-870-767-59232 | Sequence 59232, A |
| 22 | 29 | 56.9 | 154 | 4 US-09-252-991A-27472 | Sequence 27472, A |
| 23 | 29 | 56.9 | 191 | 4 US-09-252-991A-24472 | Sequence 24472, A |
| 24 | 29 | 56.9 | 213 | 4 US-09-252-991A-26933 | Sequence 26933, A |
| 25 | 29 | 56.9 | 284 | 4 US-09-252-991A-29496 | Sequence 29496, A |
| 26 | 29 | 56.9 | 284 | 4 US-09-252-991A-55993 | Sequence 25993, A |
| 27 | 29 | 56.9 | 291 | 4 US-09-252-991A-28472 | Sequence 28472, A |

| | | | | | |
|----|----|------|------|------------------------|-------------------|
| 28 | 29 | 56.9 | 302 | 4 US-09-903-814A-14 | Sequence 14, Appl |
| 29 | 29 | 56.9 | 349 | 4 US-09-489-039A-8584 | Sequence 8584, Ap |
| 30 | 29 | 56.9 | 452 | 1 US-08-336-618-23 | Sequence 23, Appl |
| 31 | 29 | 56.9 | 455 | 4 US-09-252-991A-24848 | Sequence 24848, A |
| 32 | 29 | 56.9 | 522 | 4 US-09-489-039A-10999 | Sequence 10999, A |
| 33 | 29 | 56.9 | 560 | 1 US-08-336-618-22 | Sequence 22, Appl |
| 34 | 29 | 56.9 | 611 | 4 US-09-252-991A-23379 | Sequence 23379, A |
| 35 | 29 | 56.9 | 711 | 4 US-09-252-991A-22259 | Sequence 22259, A |
| 36 | 29 | 56.9 | 6396 | 4 US-09-410-551B-72 | Sequence 72, Appl |
| 37 | 29 | 56.9 | 6396 | 4 US-09-940-316B-72 | Sequence 26906, A |
| 38 | 28 | 54.9 | 139 | 4 US-09-252-991A-26906 | Sequence 17, Appl |
| 39 | 28 | 54.9 | 146 | 3 US-09-335-409-17 | Sequence 17, Appl |
| 40 | 28 | 54.9 | 146 | 3 US-09-568-102-17 | Sequence 17, Appl |
| 41 | 28 | 54.9 | 146 | 3 US-09-567-969-17 | Sequence 17, Appl |
| 42 | 28 | 54.9 | 146 | 3 US-09-568-486-17 | Sequence 17, Appl |
| 43 | 28 | 54.9 | 146 | 3 US-09-568-486-17 | Sequence 17, Appl |
| 44 | 28 | 54.9 | 146 | 3 US-09-568-472-17 | Sequence 17, Appl |
| 45 | 28 | 54.9 | 146 | 3 US-09-567-899-17 | Sequence 17, Appl |

ALIGNMENTS

RESULT 1
US-09-878-281A-148
; Sequence 148, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Immunogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878, 281A
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 115
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-148

Query Match 68.6%; Score 35; DB 4; Length 115;
Best Local Similarity 63.6%; Pred. No. 1.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 HRTXSRXXVR 13
DB 19 HRTXSRXXVR 29
RESULT 2
US-09-252-991A-20618
; Sequence 20618, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20618
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20618

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:44:54 ; Search time 14.5191 Seconds
(without alignments)
321.196 Million cell updates/sec

Title: US-09-719-277C-4
Perfect score: 51
Sequence: 1 AAHRTXSRXXVR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_GeneSeq_23Sep04:*
1: geneeqp1980s:*
2: geneeqp1990s:*
3: geneeqp2000s:*
4: geneeqp2001s:*
5: geneeqp2002s:*
6: geneeqp2003as:*
7: geneeqp2003bs:*
8: geneeqp2004s:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|----------------|--------------------|
| 1 | 47 | 92.2 | 14 AAY44549 | Aay44549 Hepatitis |
| 2 | 47 | 92.2 | 125 6 ABU61883 | Abu61883 HCV core |
| 3 | 47 | 92.2 | 143 6 ABU61867 | Abu61867 HCV core |
| 4 | 47 | 92.2 | 143 6 ABU61873 | Abu61873 HCV core |
| 5 | 47 | 92.2 | 143 6 ABU61871 | Abu61871 HCV core |
| 6 | 47 | 92.2 | 143 6 ABU61872 | Abu61872 HCV core |
| 7 | 47 | 92.2 | 161 6 ABU61870 | Abu61870 HCV core |
| 8 | 47 | 92.2 | 198 3 AAY44558 | Aay44558 Hepatitis |
| 9 | 46 | 90.2 | 198 3 AAY44553 | Aay44553 Hepatitis |
| 10 | 45 | 88.2 | 13 3 AAY44547 | Aay44547 Hepatitis |
| 11 | 44 | 86.3 | 139 6 ABU61869 | Abu61869 HCV core |
| 12 | 43 | 84.3 | 143 6 ABU61868 | Abu61868 HCV core |
| 13 | 43 | 84.3 | 143 6 ABU61875 | Abu61875 HCV core |
| 14 | 43 | 84.3 | 160 8 ADG22790 | Adg22790 Hepatitis |
| 15 | 43 | 84.3 | 161 6 ABU61866 | Abu61866 HCV core |
| 16 | 43 | 84.3 | 161 6 ABU61864 | Abu61864 HCV core |
| 17 | 43 | 84.3 | 198 3 AAY44556 | Aay44556 Hepatitis |
| 18 | 43 | 84.3 | 198 3 AAY44554 | Aay44554 Hepatitis |
| 19 | 42 | 82.4 | 198 3 AAY44554 | Aay44554 Hepatitis |
| 20 | 40 | 78.4 | 143 6 ABU61874 | Abu61874 HCV core |
| 21 | 39 | 76.5 | 139 6 ABU61880 | Abu61880 HCV core |
| 22 | 39 | 76.5 | 154 6 ABU61878 | Abu61878 HCV core |
| 23 | 36 | 70.6 | 154 6 ABU61879 | Abu61879 HCV core |
| 24 | 36 | 70.6 | 198 3 AAY44552 | Aay44552 Hepatitis |
| 25 | 36 | 70.6 | 421 8 ADH35335 | Adh35335 EN2M prot |

| | | | | |
|----|----|------|-----------------|---------------------|
| 26 | 35 | 68.6 | 115 2 AAR63351 | Aar63351 Hepatitis |
| 27 | 35 | 68.6 | 161 6 ABU61865 | Abu61865 HCV core |
| 28 | 35 | 68.6 | 220 7 ABO71872 | Abu71872 Pseudomon |
| 29 | 35 | 68.6 | 722 7 ADB70249 | Abd70249 C. neoform |
| 30 | 35 | 68.6 | 1467 7 ABO68911 | Abu68911 Pseudomon |
| 31 | 34 | 66.7 | 121 4 AAU49040 | Aau49040 Pseudomon |
| 32 | 34 | 66.7 | 121 6 ABM45559 | Abm45559 Pseudomon |
| 33 | 34 | 66.7 | 166 3 AAG26811 | Aag26811 Zea mays |
| 34 | 34 | 66.7 | 234 8 ABO60062 | Abu60062 Human gen |
| 35 | 34 | 66.7 | 302 3 AAG44722 | Aag44722 Zea mays |
| 36 | 33 | 64.7 | 62 4 AAU43075 | Aau43075 Pseudomon |
| 37 | 33 | 64.7 | 62 6 ABM39594 | Abm39594 Pseudomon |
| 38 | 33 | 64.7 | 202 4 AAU63393 | Aau63393 Pseudomon |
| 39 | 33 | 64.7 | 202 6 ABM59912 | Abm59912 Pseudomon |
| 40 | 33 | 64.7 | 418 7 ABO68717 | Abu68717 Pseudomon |
| 41 | 33 | 64.7 | 806 4 AAU51223 | Aau51223 Pseudomon |
| 42 | 33 | 64.7 | 806 6 ABM47742 | Abm47742 Pseudomon |
| 43 | 32 | 62.7 | 66 5 ABP34536 | Abp34536 Human ORF |
| 44 | 32 | 62.7 | 214 4 ABG10896 | Abg10896 Novel hum |
| 45 | 32 | 62.7 | 444 7 ABO69021 | Abu69021 Pseudomon |

ALIGNMENTS

RESULT 1
AAY44549
ID AAY44549 standard; peptide; 14 AA.
XX
AC AAY44549;
XX
DT 04-APR-2000 (first entry)
DE Hepatitis C virus antigenic peptide ARF #2.
XX
DE Hepatitis C virus; HCV; ARF #2; alternate reading frame; hepatitis C;
XX
KW Hepatitis C virus; immunogenic; vaccine; HCV infection; antibody;
XX
KW antigenic peptide; immunogenic; vaccine; HCV infection; antibody;
XX
KW prevention; diagnosis; therapeutic target; anti-HCV therapy.
XX
OS Hepatitis C virus.
XX
PN WO9963941-A2.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-US012929.
XX
PR 09-JUN-1999; 98US-0088670P.
XX
PR 11-JUN-1999; 98US-0089138P.
XX
PA (BRAN/) BRANCH A D.
XX
PA (WALE/) WALEMSKI J L.
XX
PA (STUM/) STUMP D D.
XX
PI Branch AD, Walewski JL, Stump DD;
XX
DR WPI; 2000-126431/11.
XX
PT Novel Hepatitis C virus peptides useful in vaccine compositions, for
XX
PT diagnosing HCV infection and as therapeutic agents.
XX
PS Claim 11; Page 44; 50pp; English.
XX
CC The present sequence is a Hepatitis C virus (HCV) antigenic peptide ARF
XX #2. The novel HCV peptide is encoded by a reading frame +1 or +2 relative
XX to the standard HCV open reading frame hence not derived from the
XX standard HCV polyprotein. The peptide elicits an immune response in
XX patients infected with HCV and are produced during HCV infection. The
XX present sequence is used as an immunogen to generate antibodies against
XX HCV protein which are useful for diagnosing HCV infection. The peptide is
XX also useful in vaccine compositions for preventing HCV infection, and as
XX a target for anti-HCV therapy

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:52:49 ; Search time 13.7888 Seconds
(without alignments)
542.461 Million cell updates/sec

Title: US-09-719-277C-3
Perfect score: 56
Sequence: 1 LNLEKEXXTPTX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 40 | 71.4 | 1693 | 2 Q86H41 | Q86H41 dictyostel |
| 2 | 39 | 69.6 | 488 | 2 Q87516 | Q87516 escherichia |
| 3 | 39 | 69.6 | 488 | 2 Q8VRA7 | Q8VRA7 escherichia |
| 4 | 39 | 69.6 | 499 | 2 Q8CMD1 | Q8CMD1 escherichia |
| 5 | 39 | 69.6 | 754 | 1 ASPH_BOVIN | Q28056 bos taurus |
| 6 | 39 | 69.6 | 811 | 2 Q86I20 | Q86I20 dictyosteli |
| 7 | 37 | 66.1 | 62 | 2 Q620G8 | Q620G8 oryza sativ |
| 8 | 37 | 66.1 | 122 | 2 BAD03638 | BAD03638 oryza sat |
| 9 | 37 | 66.1 | 361 | 2 Q00687 | Q00687 hepatitis c |
| 10 | 37 | 66.1 | 409 | 2 Q94G03 | Q94G03 oryza sativ |
| 11 | 37 | 66.1 | 520 | 1 Q97Y27 | Q97Y27 sulfolobus |
| 12 | 37 | 66.1 | 520 | 1 SPRA2_HUMAN | Q9UM82 homo sapien |
| 13 | 37 | 66.1 | 562 | 2 Q8WZ19 | Q8WZ19 aspergillus |
| 14 | 36 | 64.3 | 327 | 1 Q9ET05 | Q9ET05 marmota mon |
| 15 | 36 | 64.3 | 421 | 2 Q7NDR6 | Q7NDR6 marmota mon |
| 16 | 36 | 64.3 | 421 | 2 Q910V3 | Q910V3 pseudomonas |
| 17 | 36 | 64.3 | 441 | 2 Q6Q487 | Q6Q487 aspergillus |
| 18 | 36 | 64.3 | 563 | 2 AAS68033 | AAS68033 aspergilli |
| 19 | 36 | 64.3 | 757 | 1 TNAV_HUMAN | P10636 homo sapien |
| 20 | 36 | 64.3 | 764 | 1 Q6MNP6 | Q6MNP6 bdellovibri |
| 21 | 36 | 64.3 | 764 | 1 CAE79105 | CAE79105 bdellovib |
| 22 | 36 | 64.3 | 1202 | 1 DROM_ASCIM | P23374 asacobolus i |
| 23 | 36 | 64.3 | 1111 | 1 YML1_CAEEL | P34375 caenothadi |
| 24 | 35 | 62.5 | 143 | 2 Q8NLG3 | Q8NLG3 corynebacte |
| 25 | 35 | 62.5 | 143 | 2 CAF18959 | CAF18959 corynebac |
| 26 | 35 | 62.5 | 167 | 2 Q6TKV1 | Q6TKV1 escherichia |
| 27 | 35 | 62.5 | 167 | 2 AAO96711 | AAO96711 escherich |
| 28 | 35 | 62.5 | 203 | 1 Q9X654 | Q9X654 vibrio para |
| 29 | 35 | 62.5 | 263 | 1 VGLG_ORSVW | Q86995 ovine respri |
| 30 | 35 | 62.5 | 274 | 2 Q6CFM2 | Q6CFM2 yarrowia li |

| | | | | | |
|----|----|------|------|--------------|--------------------|
| 32 | 35 | 62.5 | 297 | 2 Q922Z2 | Q922Z2 rhizobium m |
| 33 | 35 | 62.5 | 307 | 2 Q87LQ7 | Q87LQ7 vibrio para |
| 34 | 35 | 62.5 | 381 | 2 Q7VA45 | Q7VA45 prochloroco |
| 35 | 35 | 62.5 | 427 | 2 Q928J2 | Q928J2 listeria in |
| 36 | 35 | 62.5 | 499 | 2 Q72748 | Q72748 leptospira |
| 37 | 35 | 62.5 | 499 | 2 Q8F2A5 | Q8F2A5 leptospira |
| 38 | 35 | 62.5 | 499 | 2 AAS69780 | AAS69780 leptospir |
| 39 | 35 | 62.5 | 557 | 1 ESR2_ORENI | Q9YH32 oreochromis |
| 40 | 35 | 62.5 | 557 | 2 Q9W5A9 | Q9W5A9 drosophila |
| 41 | 35 | 62.5 | 592 | 2 Q9U117 | Q9U117 drosophila |
| 42 | 35 | 62.5 | 803 | 2 Q8B187 | Q8B187 mus musculu |
| 43 | 35 | 62.5 | 1222 | 2 Q7MFT6 | Q7MFT6 vibrio vuln |
| 44 | 35 | 62.5 | 1222 | 2 Q8D4B1 | Q8D4B1 vibrio vuln |
| 45 | 35 | 62.5 | 1261 | 2 Q95TV2 | Q95TV2 drosophila |

ALIGNMENTS

RESULT 1
ID Q86H41 PRELIMINARY; PRT: 1693 AA.
AC Q86H41;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypoetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Lemmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; A015598; A0053202.1; -
KW Hypoetical protein.
SQ SEQUENCE 1693 AA; 190577 MW; 25053531B44197C7 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 1693;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 1 LNLEKEXXTPT 12
Db 588 LNLEKEXXTPT 599

RESULT 2
ID Q87516 PRELIMINARY; PRT: 488 AA.
AC Q87516;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE Transposase R6.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CF1073;
RX MEDLINE=97342757; PubMed=9199454;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: October 27, 2004, 09:53:34 ; Search time 2.6 Seconds
(without alignments)
481.084 Million cell updates/sec

Title: US-09-719-277C-3
Perfect score: 56
Sequence: 1 LNLKEKPKXTPTX 13

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: 1: PIR.79.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
5: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 39 | 69.6 | 754 | 1 BABOH | peptide-aspartate |
| 2 | 37 | 66.1 | 409 | 2 D90268 | conserved hypochet |
| 3 | 36 | 64.3 | 441 | 1 QRHUT1 | microtubule-associ |
| 4 | 36 | 64.3 | 441 | 2 C83330 | hypothetical prote |
| 5 | 36 | 64.3 | 1202 | 1 S05362 | hypothetical prote |
| 6 | 35 | 62.5 | 111 | 2 S44787 | D2007.1 protein - |
| 7 | 35 | 62.5 | 263 | 2 A48732 | attachment glycopr |
| 8 | 35 | 62.5 | 263 | 2 J02388 | glycoprotein G - o |
| 9 | 35 | 62.5 | 267 | 2 B95300 | NtB, Nitrate tran |
| 10 | 35 | 62.5 | 427 | 2 AC1627 | N-acetylglucosyl-L |
| 11 | 34 | 60.7 | 133 | 2 T05440 | hypothetical prote |
| 12 | 34 | 60.7 | 147 | 2 S42552 | proline-rich prote |
| 13 | 34 | 60.7 | 181 | 2 G96715 | transcription fact |
| 14 | 34 | 60.7 | 215 | 2 S29595 | calmodulin-like pr |
| 15 | 34 | 60.7 | 216 | 2 D84841 | calmodulin-like pr |
| 16 | 34 | 60.7 | 426 | 2 AG1961 | hypothetical prote |
| 17 | 34 | 60.7 | 426 | 2 H87307 | ubiquinol-cytochro |
| 18 | 34 | 60.7 | 572 | 2 T27869 | ephinoweylin phos |
| 19 | 33 | 58.9 | 246 | 1 A46504 | chymase (EC 3.4.21 |
| 20 | 33 | 58.9 | 439 | 2 G97336 | probable S-layer p |
| 21 | 33 | 58.9 | 515 | 2 T32175 | hypothetical prote |
| 22 | 33 | 58.9 | 516 | 2 B96839 | hypothetical prote |
| 23 | 33 | 58.9 | 613 | 2 S19676 | 68k protein - phag |
| 24 | 33 | 58.9 | 670 | 2 S67383 | probable signal tr |
| 25 | 33 | 58.9 | 1071 | 2 E85343 | hypothetical prote |
| 26 | 32 | 57.1 | 128 | 2 S77680 | streptokinase A (B |
| 27 | 32 | 57.1 | 128 | 2 S77680 | streptokinase A (B |
| 28 | 32 | 57.1 | 128 | 2 S77688 | probable chaperone |
| 29 | 32 | 57.1 | 209 | 2 D90788 | |

| | | | | | |
|----|----|------|-----|----------|---------------------|
| 30 | 32 | 57.1 | 229 | 2 E85648 | probable chaperone |
| 31 | 32 | 57.1 | 244 | 2 S36703 | gene 8 protein - e |
| 32 | 32 | 57.1 | 248 | 2 PQ0769 | glycoprotein G - b |
| 33 | 32 | 57.1 | 250 | 2 PQ0768 | glycoprotein G - b |
| 34 | 32 | 57.1 | 257 | 1 MEN2BR | major surface glyco |
| 35 | 32 | 57.1 | 263 | 2 JQ2284 | glycoprotein G - b |
| 36 | 32 | 57.1 | 267 | 2 H97299 | prob (imported) - |
| 37 | 32 | 57.1 | 329 | 2 E71941 | probable ribosomal |
| 38 | 32 | 57.1 | 333 | 1 D64653 | ribosomal protein |
| 39 | 32 | 57.1 | 349 | 2 D64134 | peptide transport |
| 40 | 32 | 57.1 | 371 | 2 A71683 | hypothetical prote |
| 41 | 32 | 57.1 | 401 | 2 G84160 | glucose-1-phosphat |
| 42 | 32 | 57.1 | 412 | 2 T46104 | hypothetical prote |
| 43 | 32 | 57.1 | 416 | 2 A43561 | homeotic protein m |
| 44 | 32 | 57.1 | 423 | 2 D96552 | unknown protein, 7 |
| 45 | 32 | 57.1 | 425 | 2 AH0978 | probable membrane |

ALIGNMENTS

RESULT 1
BABOH
peptide-aspartate beta-dioxygenase (EC 1.14.11.16) - bovine
N:Alternate names: aspartyl (asparaginyl) beta-hydroxylase
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1993 #sequence revision 10-Feb-1995 #text change 09-Jul-2004
C:Accession: A42969; A39470; B39470; S27948
R:Jia, S.; Vandusen, W.J.; Diethl, R.E.; Kohl, N.E.; Dixon, R.A.; Elliston, K.O.; Stern, J. Biol. Chem. 267, 14322-14327, 1992
J. Title: Bovine liver aspartyl beta-hydroxylase. Purification and characterization.
A:Reference number: A39470; MUID:91310689; PMID:1856229
A:Accession: A39470
A:Reference type: mRNA
A:Molecule type: protein
A:Residues: 1-754 <JIA>
A:Cross-references: UNIPROT:Q28056; EMBL:M91213; NID:g162693; PIDN:AAA03563.1; PID:g162
A:Experimental source: brain
A:Note: Sequence extracted from NCBI backbone (NCBI:108534)
R:Wang, Q.; Vandusen, W.J.; Petrovski, C.U.; Garsky, V.M.; Stern, A.M.; Friedman, P.A.
J. Biol. Chem. 266, 14004-14010, 1991
J. Title: Bovine liver aspartyl beta-hydroxylase. Purification and characterization.
A:Reference number: A39470; MUID:91310689; PMID:1856229
A:Accession: A39470
A:Reference type: protein
A:Molecule type: protein
A:Residues: 289-328 <MAN>
A:Accession: B39470
A:Molecule type: protein
A:Residues: 615, 'X', 617-630, 'XX', 633-634, 'X', 636, 'XX', 639-641 <MA2>
A:Accession: C39470
A:Molecule type: protein
A:Residues: 311-347, 'X', 349, 'X', 351-373, 'X', 375-379, 'X', 381-382 <MA3>
A:Comment: This enzyme uses ferrous iron as a cofactor, and while beta-hydroxylating C
C:Comment: Aspartic acid and asparagine residues in the BGP homology domain of certain
C:Superfamily: peptide-aspartate beta-dioxygenase; tetratricopeptide repeat homology
C:Keywords: glycoprotein; oxidoreductase; transmembrane protein
F:2-56/Domain: intracellular #status predicted <TMC>
F:57-78/Domain: transmembrane #status predicted <TMC>
F:289-754/Product: peptide-aspartate beta-dioxygenase, 56k form #status predicted <56k
F:311-754/Product: peptide-aspartate beta-dioxygenase, 52k form #status predicted <52k
F:337-370/Domain: tetratricopeptide repeat homology <TT1>
F:371-404/Domain: tetratricopeptide repeat homology (covalent) #status predicted
F:113,96,466,702/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.6%; Score 39; DB 1; Length 754;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservativity 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 LNLKEKPKXTPT 12
DB 125 LNLKEKPKXTPT 136
RESULT 2

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:56:09 ; Search time 10.1663 Seconds
(without alignments)
414.586 Million cell updates/sec

Title: US-09-719-277C-3
Perfect score: 56
Sequence: 1 INLKEKPTXTPTX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------|--------------------|
| 1 | 51 | 91.1 | 13 | US-10-601-020-3 | Sequence 3, Appl1 |
| 2 | 51 | 91.1 | 14 | US-10-601-020-5 | Sequence 5, Appl1 |
| 3 | 51 | 91.1 | 192 | US-10-601-020-17 | Sequence 17, Appl1 |
| 4 | 51 | 91.1 | 194 | US-10-601-020-14 | Sequence 14, Appl1 |
| 5 | 51 | 91.1 | 195 | US-10-601-020-9 | Sequence 9, Appl1 |
| 6 | 51 | 91.1 | 195 | US-10-601-020-13 | Sequence 13, Appl1 |
| 7 | 51 | 91.1 | 195 | US-10-601-020-16 | Sequence 16, Appl1 |
| 8 | 51 | 91.1 | 196 | US-10-601-020-2 | Sequence 2, Appl1 |
| 9 | 51 | 91.1 | 196 | US-10-601-020-10 | Sequence 10, Appl1 |
| 10 | 47 | 83.9 | 192 | US-10-601-020-11 | Sequence 11, Appl1 |
| 11 | 43 | 76.8 | 197 | US-10-601-020-12 | Sequence 12, Appl1 |
| 12 | 39 | 69.6 | 160 | US-10-150-283-1 | Sequence 14, Appl1 |
| 13 | 38 | 67.9 | 59 | US-10-012-542-481 | Sequence 481, App |

| | | | | | | |
|----|----|------|------|----|----------------------|----------------------|
| 14 | 38 | 67.9 | 59 | 14 | US-10-115-123-481 | Sequence 481, App |
| 15 | 38 | 67.9 | 114 | 14 | US-10-012-542-479 | Sequence 479, App |
| 16 | 38 | 67.9 | 114 | 14 | US-10-115-123-479 | Sequence 479, App |
| 17 | 37 | 66.1 | 74 | 15 | US-10-424-599-275115 | Sequence 275115, App |
| 18 | 37 | 66.1 | 117 | 16 | US-10-437-963-161684 | Sequence 161684, App |
| 19 | 37 | 66.1 | 194 | 16 | US-10-601-020-15 | Sequence 15, Appl1 |
| 20 | 37 | 66.1 | 520 | 16 | US-10-408-765A-570 | Sequence 570, App |
| 21 | 37 | 66.1 | 768 | 16 | US-10-437-963-118237 | Sequence 118237, App |
| 22 | 37 | 66.1 | 2396 | 16 | US-10-437-963-122555 | Sequence 122555, App |
| 23 | 36 | 64.3 | 14 | 14 | US-10-256-865-15 | Sequence 15, Appl1 |
| 24 | 36 | 64.3 | 118 | 16 | US-10-767-701-36004 | Sequence 36004, App |
| 25 | 36 | 64.3 | 381 | 15 | US-10-470-928-4 | Sequence 4, Appl1 |
| 26 | 36 | 64.3 | 400 | 15 | US-10-470-928-2 | Sequence 2, Appl1 |
| 27 | 36 | 64.3 | 410 | 15 | US-10-470-928-5 | Sequence 5, Appl1 |
| 28 | 36 | 64.3 | 412 | 15 | US-10-470-928-9 | Sequence 9, Appl1 |
| 29 | 36 | 64.3 | 441 | 9 | US-09-142-613-1 | Sequence 1, Appl1 |
| 30 | 36 | 64.3 | 441 | 10 | US-09-847-586-79 | Sequence 79, Appl1 |
| 31 | 36 | 64.3 | 441 | 13 | US-10-076-708-2 | Sequence 2, Appl1 |
| 32 | 36 | 64.3 | 441 | 13 | US-10-107-181-5 | Sequence 5, Appl1 |
| 33 | 36 | 64.3 | 441 | 13 | US-10-066-810-1 | Sequence 1, Appl1 |
| 34 | 36 | 64.3 | 441 | 15 | US-10-451-367-2 | Sequence 2, Appl1 |
| 35 | 36 | 64.3 | 441 | 15 | US-10-470-928-6 | Sequence 6, Appl1 |
| 36 | 36 | 64.3 | 441 | 16 | US-10-451-782-2 | Sequence 2, Appl1 |
| 37 | 36 | 64.3 | 441 | 17 | US-10-402-420-3 | Sequence 3, Appl1 |
| 38 | 36 | 64.3 | 447 | 16 | US-10-693-999-1 | Sequence 1, Appl1 |
| 39 | 36 | 64.3 | 554 | 14 | US-10-369-493-13265 | Sequence 13265, App |
| 40 | 36 | 64.3 | 758 | 9 | US-09-904-987-5 | Sequence 5, Appl1 |
| 41 | 36 | 64.3 | 758 | 16 | US-10-648-593-242 | Sequence 242, App |
| 42 | 35 | 62.5 | 143 | 9 | US-09-738-626-6841 | Sequence 6841, App |
| 43 | 35 | 62.5 | 220 | 15 | US-10-424-589-22351 | Sequence 22351, App |
| 44 | 35 | 62.5 | 1300 | 16 | US-10-408-765A-867 | Sequence 867, App |
| 45 | 34 | 60.7 | 96 | 15 | US-10-424-599-214255 | Sequence 214255, App |

ALIGNMENTS

RESULT 1
US-10-601-020-3
; Sequence 3, Application US/10601020
; Publication No. US20040156862A1
; GENERAL INFORMATION:
; APPLICANT: Branch, Andrea D.
; APPLICANT: Walewski, Jose L.
; APPLICANT: Stump, Decherd D.
; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
; FILE REFERENCE: RIT-003CRUSCN
; CURRENT APPLICATION NUMBER: US/10/601,020
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/719277
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/088670
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: US 60/089138
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: PCT/US99/12929
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: Variant
; LOCATION: 8
; OTHER INFORMATION: Xaa = Asn or Lys
; FEATURE:
; NAME/KEY: Variant
; LOCATION: 9
; OTHER INFORMATION: Xaa = Val or Glu
; FEATURE:
; NAME/KEY: Variant

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:53:49 ; Search time 3.4764 Seconds
(without alignments)
247.996 Million cell updates/sec

Title: US-09-719-277C-3
Perfect score: 56
Sequence: 1 LNLKEXPTPTX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 38 | 67.9 | 59 | 4 | US-09-461-325-481 |
| 2 | 38 | 67.9 | 59 | 4 | US-10-012-542-481 |
| 3 | 38 | 67.9 | 59 | 4 | US-10-115-123-481 |
| 4 | 38 | 67.9 | 114 | 4 | US-09-461-325-479 |
| 5 | 38 | 67.9 | 114 | 4 | US-10-012-542-479 |
| 6 | 38 | 67.9 | 114 | 4 | US-10-115-123-479 |
| 7 | 36 | 64.3 | 14 | 4 | US-09-252-404A-15 |
| 8 | 36 | 64.3 | 441 | 3 | US-08-244-603A-1 |
| 9 | 36 | 64.3 | 441 | 3 | US-08-913-915-5 |
| 10 | 36 | 64.3 | 441 | 4 | US-09-035-708A-1 |
| 11 | 36 | 64.3 | 555 | 4 | US-09-252-991A-31656 |
| 12 | 36 | 64.3 | 758 | 4 | US-09-904-987-5 |
| 13 | 35 | 62.5 | 375 | 4 | US-09-489-039A-9807 |
| 14 | 35 | 62.5 | 577 | 4 | US-09-248-796A-18108 |
| 15 | 35 | 62.5 | 981 | 4 | US-09-252-991A-24231 |
| 16 | 34 | 60.7 | 147 | 2 | US-08-530-165-3 |
| 17 | 34 | 60.7 | 147 | 2 | US-08-530-165-3 |
| 18 | 33 | 58.9 | 147 | 4 | US-09-252-991A-24180 |
| 19 | 33 | 58.9 | 240 | 3 | US-09-329-418-7 |
| 20 | 33 | 58.9 | 240 | 3 | US-09-531-914-7 |
| 21 | 33 | 58.9 | 249 | 4 | US-09-673-395A-338 |
| 22 | 33 | 58.9 | 262 | 4 | US-09-673-395A-597 |
| 23 | 33 | 58.9 | 292 | 4 | US-09-107-532A-7226 |
| 24 | 33 | 58.9 | 404 | 4 | US-09-445-724B-4 |
| 25 | 33 | 58.9 | 404 | 4 | US-09-445-724B-8 |
| 26 | 33 | 58.9 | 420 | 3 | US-09-329-418-8 |
| 27 | 33 | 58.9 | 420 | 3 | US-09-531-914-8 |

| | | | | | | |
|----|----|------|------|---|---------------------|-------------------|
| 28 | 33 | 58.9 | 455 | 4 | US-09-328-352-5626 | Sequence 5626, Ap |
| 29 | 33 | 58.9 | 497 | 4 | US-09-345-473B-8 | Sequence 8, Appl |
| 30 | 33 | 58.9 | 518 | 3 | US-09-329-418-3 | Sequence 3, Appl |
| 31 | 33 | 58.9 | 518 | 3 | US-09-329-418-4 | Sequence 4, Appl |
| 32 | 33 | 58.9 | 518 | 3 | US-09-329-418-5 | Sequence 5, Appl |
| 33 | 33 | 58.9 | 518 | 3 | US-09-329-418-9 | Sequence 9, Appl |
| 34 | 33 | 58.9 | 518 | 3 | US-09-531-914-3 | Sequence 3, Appl |
| 35 | 33 | 58.9 | 518 | 3 | US-09-531-914-4 | Sequence 4, Appl |
| 36 | 33 | 58.9 | 518 | 3 | US-09-531-914-5 | Sequence 5, Appl |
| 37 | 33 | 58.9 | 519 | 3 | US-09-531-914-9 | Sequence 9, Appl |
| 38 | 33 | 58.9 | 519 | 3 | US-08-997-445D-2 | Sequence 2, Appl |
| 39 | 33 | 58.9 | 541 | 4 | US-09-270-767-59648 | Sequence 59648, A |
| 40 | 33 | 58.9 | 629 | 4 | US-09-270-767-44226 | Sequence 44226, A |
| 41 | 33 | 58.9 | 1122 | 4 | US-09-489-039A-8554 | Sequence 8554, Ap |
| 42 | 32 | 57.1 | 126 | 4 | US-09-270-767-39597 | Sequence 39597, A |
| 43 | 32 | 57.1 | 126 | 4 | US-09-270-767-54814 | Sequence 54814, A |
| 44 | 32 | 57.1 | 223 | 4 | US-09-328-352-8104 | Sequence 8104, Ap |
| 45 | 32 | 57.1 | 257 | 4 | US-09-567-458A-2 | Sequence 2, Appl |

ALIGNMENTS

RESULT 1
US-09-461-325-481
Sequence 481, Application US/09461325A
Patent No. 6475753
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P2029P1
CURRENT APPLICATION NUMBER: US/09/461,325A
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: PCT/US99/13418
EARLIER FILING DATE: 1999-06-15
EARLIER APPLICATION NUMBER: 60/089, 507
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089, 508
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089, 509
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089, 510
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/090, 112
EARLIER FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: 60/090, 113
EARLIER FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 481
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-09-461-325-481

Query Match 67.9% Score 38; DB 4; Length 59;
Best Local Similarity 66.7% Pred. No. 3;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LNLKEXPTPTX 12
DB 22 LNLKEXPTPTX 33

RESULT 2
US-10-012-542-481
Sequence 481, Application US/10012542
Patent No. 6627741
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P2029P1
CURRENT APPLICATION NUMBER: US/10/012,542

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:44:54 ; Search time 14.5191 Seconds
(without alignments)
321.196 Million cell updates/sec

Title: US-09-719-277C-3
Perfect score: 56
Sequence: 1 INLKEKPYXPTX 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|-----|--------------------|
| 1 | 51 | 91.1 | 13 | AAV44546 Hepatitis |
| 2 | 51 | 91.1 | 14 | AAV44548 Hepatitis |
| 3 | 51 | 91.1 | 198 | AAV44555 Hepatitis |
| 4 | 51 | 91.1 | 198 | AAV44557 Hepatitis |
| 5 | 51 | 91.1 | 198 | AAV44554 Hepatitis |
| 6 | 51 | 91.1 | 198 | AAV44545 Hepatitis |
| 7 | 51 | 91.1 | 198 | AAV44558 Hepatitis |
| 8 | 47 | 83.9 | 198 | AAV44552 Hepatitis |
| 9 | 43 | 76.8 | 198 | AAV44553 Hepatitis |
| 10 | 39 | 69.6 | 160 | ADG22790 Hepatitis |
| 11 | 38 | 67.9 | 59 | AAV86539 Human gen |
| 12 | 38 | 67.9 | 59 | ABOS3705 Novel hum |
| 13 | 38 | 67.9 | 114 | AAV86537 Human gen |
| 14 | 38 | 67.9 | 114 | ABOS3703 Novel hum |
| 15 | 37 | 66.1 | 198 | AAV44556 Hepatitis |
| 16 | 37 | 66.1 | 520 | ADJ68764 Human hea |
| 17 | 36 | 64.3 | 12 | AAV37560 Phosphory |
| 18 | 36 | 64.3 | 12 | AAV38241 Alzheimer |
| 19 | 36 | 64.3 | 14 | AAV21957 Phosphory |
| 20 | 36 | 64.3 | 215 | ADL22478 Human tau |
| 21 | 36 | 64.3 | 381 | AAV27915 Human tau |
| 22 | 36 | 64.3 | 390 | AAV05283 Truncated |
| 23 | 36 | 64.3 | 400 | AAV27913 Human C-c |
| 24 | 36 | 64.3 | 410 | AAV27916 Human tau |
| 25 | 36 | 64.3 | 412 | AAV27920 Human tau |

| | | | | | |
|----|----|------|-----|---|---------------------|
| 26 | 36 | 64.3 | 441 | 2 | AAV58810 Human tau |
| 27 | 36 | 64.3 | 441 | 2 | AAV05282 Human tau |
| 28 | 36 | 64.3 | 441 | 2 | AAV34856 Human tau |
| 29 | 36 | 64.3 | 441 | 3 | AAV81386 Human tau |
| 30 | 36 | 64.3 | 441 | 5 | ABV98513 Human tau |
| 31 | 36 | 64.3 | 441 | 5 | ABV52186 Human tau |
| 32 | 36 | 64.3 | 441 | 5 | ABV05642 Mutant hu |
| 33 | 36 | 64.3 | 441 | 5 | ABV31741 Human tau |
| 34 | 36 | 64.3 | 441 | 5 | AAV27917 Human tau |
| 35 | 36 | 64.3 | 441 | 5 | ABV09398 Human tau |
| 36 | 36 | 64.3 | 441 | 5 | AAV14577 Human Cys |
| 37 | 36 | 64.3 | 441 | 5 | ADV52325 W1d-type |
| 38 | 36 | 64.3 | 441 | 7 | ADK19389 human cen |
| 39 | 36 | 64.3 | 447 | 8 | ADV31115 N-end ru1 |
| 40 | 36 | 64.3 | 555 | 7 | ABV082910 Pseudomon |
| 41 | 36 | 64.3 | 758 | 5 | ABV178011 Amino aci |
| 42 | 36 | 64.3 | 758 | 8 | ADL61318 Human tyr |
| 43 | 35 | 62.5 | 143 | 4 | AAV93087 C glutam |
| 44 | 35 | 62.5 | 263 | 8 | ADL90097 Major sur |
| 45 | 35 | 62.5 | 375 | 7 | ABV63290 Klebsiell |

ALIGNMENTS

| | | |
|----------|---|------------------------------------|
| RESULT 1 | AAV44546 | AAV44546 standard; peptide; 13 AA. |
| ID | AAV44546 | |
| XX | AAV44546 | |
| AC | AAV44546 | |
| DT | 04-APR-2000 | (first entry) |
| DE | Hepatitis C virus antigenic peptide ARF #1 consensus sequence. | |
| XX | Hepatitis C virus; HCV; ARF #1 antigenic peptide consensus sequence; | |
| KW | alternate reading frame; hepatitis C; immunogenic; vaccine; prevention; | |
| KW | diagnosis; therapeutic target; anti-HCV therapy; HCV infection; antibody. | |
| XX | Hepatitis C virus. | |
| OS | Hepatitis C virus. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | Misc-difference 8 | /label= Asn, Lys |
| FT | Misc-difference 9 | /label= Val, Glu |
| FT | Misc-difference 13 | /label= Ala, Val |
| XX | | |
| PN | WO963941-A2. | |
| XX | 16-DEC-1999. | |
| PD | | |
| XX | 09-JUN-1999; | 99WO-US012929. |
| PF | | |
| XX | 09-JUN-1998; | 98US-0088670P. |
| PR | 11-JUN-1998; | 98US-0089138P. |
| XX | | |
| PA | (BRAN/) BRANCH A D. | |
| PA | (WALE/) WALEMSKI J L. | |
| PA | (STUM/) STUMP D D. | |
| PI | Branch AD, Walewski JL, Stump DD; | |
| DR | WPI; 2000-126431/11. | |
| XX | | |
| XX | Novel Hepatitis C virus peptides useful in vaccine compositions, for | |
| PT | diagnosing HCV infection and as therapeutic agents. | |
| XX | | |
| PS | Claim 11; Page 43; 50pp; English. | |
| XX | | |
| CC | The present peptide is a consensus sequence of Hepatitis C virus (HCV) | |
| CC | antigenic peptide ARF #1. The novel HCV peptide is encoded by a reading | |

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:56:09 ; Search time 153.276 Seconds
(without alignments)
414.586 Million cell updates/sec

Title: US-09-719-277C-2
Perfect score: 1039
Sequence: 1 ARIINLKEKNVPTVAHRT.....WPCSLALCPLOPKCAIRPG 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
17: /cgn2_6/prodata/1/pubppaa/US10_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubppaa/US11_NEW_PUB.pep.*
19: /cgn2_6/prodata/1/pubppaa/US60_NEW_PUB.pep.*
20: /cgn2_6/prodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match Length | ID | Description |
|------------|-------|-------|--------------|-------------------|--------------------|
| 1 | 1039 | 100.0 | 196 | US-10-601-020-2 | Sequence 2, Appl1 |
| 2 | 1039 | 100.0 | 196 | US-10-601-020-10 | Sequence 10, Appl1 |
| 3 | 804 | 77.4 | 160 | US-10-150-283-1 | Sequence 1, Appl1 |
| 4 | 787 | 75.7 | 161 | US-09-736-959A-2 | Sequence 2, Appl1 |
| 5 | 773 | 74.4 | 161 | US-09-736-959A-1 | Sequence 1, Appl1 |
| 6 | 773 | 74.4 | 161 | US-09-736-959A-32 | Sequence 32, Appl1 |
| 7 | 766 | 73.7 | 161 | US-09-736-959A-3 | Sequence 3, Appl1 |
| 8 | 745.5 | 71.8 | 195 | US-10-601-020-9 | Sequence 9, Appl1 |
| 9 | 715 | 68.8 | 192 | US-10-601-020-17 | Sequence 17, Appl1 |
| 10 | 649.5 | 62.5 | 195 | US-10-601-020-13 | Sequence 13, Appl1 |
| 11 | 622 | 59.9 | 161 | US-09-736-959A-7 | Sequence 7, Appl1 |
| 12 | 561.5 | 54.0 | 195 | US-10-601-020-16 | Sequence 16, Appl1 |
| 13 | 555 | 53.4 | 143 | US-09-736-959A-5 | Sequence 5, Appl1 |

| | | | | | | |
|----|-------|------|-------|----|----------------------|--------------------|
| 14 | 551 | 53.0 | 143 | 9 | US-09-736-959A-9 | Sequence 9, Appl1 |
| 15 | 550 | 52.9 | 143 | 9 | US-09-736-959A-8 | Sequence 8, Appl1 |
| 16 | 545 | 52.5 | 143 | 9 | US-09-736-959A-12 | Sequence 12, Appl1 |
| 17 | 543 | 52.3 | 143 | 9 | US-09-736-959A-4 | Sequence 4, Appl1 |
| 18 | 543 | 52.3 | 143 | 9 | US-09-736-959A-10 | Sequence 10, Appl1 |
| 19 | 543 | 52.3 | 143 | 9 | US-09-736-959A-11 | Sequence 11, Appl1 |
| 20 | 536 | 51.6 | 192 | 16 | US-10-601-020-11 | Sequence 15, Appl1 |
| 21 | 532 | 51.2 | 194 | 16 | US-10-601-020-15 | Sequence 16, Appl1 |
| 22 | 493 | 47.4 | 154 | 9 | US-09-736-959A-16 | Sequence 12, Appl1 |
| 23 | 491 | 47.3 | 197 | 16 | US-10-601-020-12 | Sequence 14, Appl1 |
| 24 | 470 | 45.2 | 194 | 16 | US-10-601-020-14 | Sequence 6, Appl1 |
| 25 | 431 | 41.5 | 139 | 9 | US-09-736-959A-6 | Sequence 17, Appl1 |
| 26 | 431 | 41.5 | 139 | 9 | US-09-736-959A-17 | Sequence 15, Appl1 |
| 27 | 424 | 40.8 | 154 | 9 | US-09-736-959A-15 | Sequence 20, Appl1 |
| 28 | 421 | 40.5 | 125 | 9 | US-09-736-959A-20 | Sequence 19, Appl1 |
| 29 | 365 | 35.1 | 125 | 9 | US-09-736-959A-14 | Sequence 19, Appl1 |
| 30 | 362 | 34.8 | 125 | 9 | US-09-736-959A-21 | Sequence 21, Appl1 |
| 31 | 360 | 34.6 | 125 | 9 | US-09-736-959A-18 | Sequence 18, Appl1 |
| 32 | 351 | 33.8 | 125 | 9 | US-09-736-959A-18 | Sequence 14, Appl1 |
| 33 | 322 | 31.0 | 115 | 10 | US-09-873-224-148 | Sequence 13, Appl1 |
| 34 | 314 | 30.2 | 125 | 9 | US-09-736-959A-13 | Sequence 8, Appl1 |
| 35 | 116.5 | 11.2 | 19608 | 15 | US-10-084-846A-8 | Sequence 3324, Ap |
| 36 | 100.5 | 9.7 | 602 | 14 | US-10-104-047-3324 | Sequence 3, Appl1 |
| 37 | 98.5 | 9.5 | 19695 | 15 | US-10-084-846A-3 | Sequence 13190, A |
| 38 | 98 | 9.4 | 705 | 14 | US-10-156-761-13190 | Sequence 114227, A |
| 39 | 96.5 | 9.3 | 205 | 16 | US-10-437-963-114227 | Sequence 7, Appl1 |
| 40 | 96.5 | 9.3 | 19652 | 15 | US-10-084-846A-7 | Sequence 4069, Ap |
| 41 | 95.5 | 9.2 | 220 | 15 | US-10-108-260A-4069 | Sequence 183009, A |
| 42 | 94 | 9.0 | 254 | 16 | US-10-437-963-183009 | Sequence 2648, Ap |
| 43 | 94 | 9.0 | 291 | 14 | US-10-094-749-2648 | Sequence 5, Appl1 |
| 44 | 93.5 | 9.0 | 19723 | 15 | US-10-084-846A-5 | Sequence 164016, A |
| 45 | 93 | 9.0 | 2246 | 16 | US-10-437-963-164016 | |

ALIGNMENTS

RESULT 1
US-10-601-020-2
Sequence 2, Application US/10601020
Publication No. US20040156862A1
GENERAL INFORMATION:
APPLICANT: Branch, Andrea D.
APPLICANT: Malewski, Jose L.
APPLICANT: Stump, Dechard D.
TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
FILE REFERENCE: RII-003CPUSCN
CURRENT APPLICATION NUMBER: US/10/601,020
CURRENT FILING DATE: 2003-06-20
PRIOR APPLICATION NUMBER: US 09/719277
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/088670
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: US 60/089138
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: PCT/US99/12929
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 196
TYPE: PRT
ORGANISM: Hepatitis C virus
US-10-601-020-2

Query Match 100.0%; Score 1039; DB 16; Length 196;
Best Local Similarity 100.0%; Pred. No. 6.1e-84;
Matches 196; Conserved 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ARIINLKEKNVPTVAHRTSSRVAVRSLVEFTCCAGALDWVCARGRLPGGRNLEVD 60
DB 1 ARIINLKEKNVPTVAHRTSSRVAVRSLVEFTCCAGALDWVCARGRLPGGRNLEVD 60

OM protein - protein search, using SW model

Run on: October 27, 2004, 09:44:54 ; Search time 218.903 Seconds
(without alignments)
321.196 Million cell updates/sec

Title: US-09-719-277C-2
Perfect score: 1039
Sequence: 1 ARINLKEKPVPTVAHRT.....WPCSLALCPLOPTKCAIRPG 196

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 1018 | 98.0 | 198 | 3 | AAV44545 Hepatitis |
| 2 | 804 | 77.4 | 160 | 8 | ADG22790 Hepatitis |
| 3 | 787 | 75.7 | 161 | 6 | ABU61865 HCV core |
| 4 | 773 | 74.4 | 161 | 6 | ABU61864 HCV core |
| 5 | 766 | 73.7 | 161 | 6 | ABU61866 HCV core |
| 6 | 727 | 70.0 | 198 | 3 | AAV44558 Hepatitis |
| 7 | 648 | 62.4 | 198 | 3 | AAV44554 Hepatitis |
| 8 | 622 | 59.9 | 161 | 6 | ABU61870 HCV core |
| 9 | 569 | 54.8 | 198 | 3 | AAV44557 Hepatitis |
| 10 | 555 | 53.4 | 143 | 6 | ABU61868 HCV core |
| 11 | 554.5 | 53.4 | 198 | 3 | AAV44552 Hepatitis |
| 12 | 551 | 53.0 | 143 | 6 | ABU61872 HCV core |
| 13 | 550 | 52.9 | 143 | 6 | ABU61871 HCV core |
| 14 | 545 | 52.5 | 143 | 6 | ABU61875 HCV core |
| 15 | 543 | 52.3 | 143 | 6 | ABU61867 HCV core |
| 16 | 543 | 52.3 | 143 | 6 | ABU61873 HCV core |
| 17 | 543 | 52.3 | 143 | 6 | ABU61874 HCV core |
| 18 | 512 | 49.3 | 198 | 3 | AAV44556 Hepatitis |
| 19 | 495.5 | 47.7 | 198 | 3 | AAV44555 Hepatitis |
| 20 | 493 | 47.4 | 154 | 6 | ABU61879 HCV core |
| 21 | 483.5 | 46.5 | 198 | 3 | AAV44553 Hepatitis |
| 22 | 431 | 41.5 | 139 | 6 | ABU61880 HCV core |
| 23 | 424 | 41.5 | 139 | 6 | ABU61869 HCV core |
| 24 | 424 | 40.8 | 154 | 6 | ABU61878 HCV core |
| 25 | 421 | 40.5 | 125 | 6 | ABU61883 HCV core |

| | | | | | |
|----|-------|------|------|---|--------------------|
| 26 | 400 | 38.5 | 76 | 8 | ADK02062 Hepatitis |
| 27 | 396 | 38.1 | 76 | 8 | ADK01350 Hepatitis |
| 28 | 365 | 35.1 | 125 | 6 | ABU61877 HCV core |
| 29 | 362 | 34.8 | 125 | 6 | ABU61882 HCV core |
| 30 | 360 | 34.6 | 125 | 6 | ABU61884 HCV core |
| 31 | 351 | 33.8 | 125 | 6 | ABU61881 HCV core |
| 32 | 322 | 31.0 | 115 | 2 | AAR63351 Hepatitis |
| 33 | 314 | 30.2 | 125 | 6 | ABU61876 HCV core |
| 34 | 287 | 27.6 | 91 | 8 | ADK01937 Hepatitis |
| 35 | 251 | 24.2 | 58 | 8 | ADK01459 Hepatitis |
| 36 | 202 | 19.4 | 69 | 8 | ADK01829 Hepatitis |
| 37 | 190 | 18.3 | 76 | 8 | ADK01709 Hepatitis |
| 38 | 110.5 | 10.6 | 1938 | 6 | ABP76682 Streptomy |
| 39 | 109 | 10.5 | 21 | 8 | ADK02063 Hepatitis |
| 40 | 109 | 10.5 | 21 | 8 | ADK01351 Hepatitis |
| 41 | 109 | 10.5 | 236 | 7 | ABO78872 Pseudomon |
| 42 | 104.5 | 10.1 | 328 | 7 | ABO68983 Pseudomon |
| 43 | 100.5 | 9.7 | 602 | 7 | ADB65170 Human pro |
| 44 | 100 | 9.6 | 303 | 7 | ABO71053 Pseudomon |
| 45 | 100 | 9.6 | 709 | 7 | ABO78559 Pseudomon |

ALIGNMENTS

RESULT 1
ID AAV44545 standard; protein; 198 AA.
AC AAV44545;
DT 04-APR-2000 (first entry)
XX
DE Hepatitis C virus protein encoded by DNA clone AF011751.
XX
KW Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine;
KW prevention; diagnosis; therapeutic target; anti-HCV therapy;
KW HCV infection; DNA clone AF011751.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 146 /note="Encoded by CNG"
FT Misc-difference 161 /label="unknown"
FT FT /note="Encoded by TGA"
FT Misc-difference 184 /label="unknown"
FT FT /note="Encoded by TGA"
XX
MO963941-A2.
XX
PN 16-DEC-1999.
XX
PD 09-JUN-1999; 99WC-US012929.
XX
PF 09-JUN-1998; 98US-0086670P.
XX
PR 11-JUN-1998; 98US-0089138P.
XX
PA (BRAN/) BRANCH A D.
PA (WALE/) WALEWSKI J L.
XX (STUM/) STUMP D D.
XX
PI Branch AD, Walewski JL, Stump DD;
XX
XX WPI, 2000-126431/11.
XX DR N-PSDB; AAZ29897.
XX
XX Novel Hepatitis C virus peptides useful in vaccine compositions, for
XX diagnosing HCV infection and as therapeutic agents.
XX
XX Claim 1, Page 39-41; 50pp; English.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:53:49 ; Search time 52.4135 seconds
(without alignments)
247.996 Million cell updates/sec

Title: US-09-719-277C-2

Perfect score: 1039

Sequence: 1 ARIINLKEKPNVTPTVAHRT.....WPCSLALCPLOPTKCAIRPG 196

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA:*
1: /cgn2_6/p/codata/1/1aa/5B_COMB.pep:*
2: /cgn2_6/p/codata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/p/codata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/p/codata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/p/codata/1/1aa/6C_COMB.pep:*
6: /cgn2_6/p/codata/1/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 322 | 31.0 | 115 | 4 | US-09-878-281A-148 Sequence 148, App |
| 2 | 109 | 10.5 | 236 | 4 | US-09-252-991A-27618 Sequence 27618, A |
| 3 | 104.5 | 10.1 | 328 | 4 | US-09-252-991A-17729 Sequence 17729, A |
| 4 | 100 | 9.6 | 303 | 4 | US-09-252-991A-17729 Sequence 17729, A |
| 5 | 100 | 9.6 | 709 | 4 | US-09-252-991A-27305 Sequence 27305, A |
| 6 | 99.5 | 9.6 | 263 | 4 | US-09-252-991A-16906 Sequence 16906, A |
| 7 | 98.5 | 9.5 | 1498 | 4 | US-09-252-991A-31234 Sequence 31234, A |
| 8 | 96.5 | 9.3 | 165 | 4 | US-09-252-991A-28858 Sequence 28858, A |
| 9 | 96.5 | 9.3 | 573 | 4 | US-09-252-991A-20056 Sequence 20056, A |
| 10 | 96 | 9.2 | 202 | 4 | US-09-252-991A-24790 Sequence 24790, A |
| 11 | 96 | 9.2 | 242 | 4 | US-09-252-991A-31425 Sequence 31425, A |
| 12 | 95.5 | 9.2 | 194 | 4 | US-09-252-991A-28860 Sequence 28860, A |
| 13 | 95.5 | 9.2 | 231 | 4 | US-09-252-991A-22801 Sequence 22801, A |
| 14 | 95 | 9.1 | 1027 | 4 | US-09-252-991A-26216 Sequence 26216, A |
| 15 | 94.5 | 9.1 | 171 | 4 | US-09-252-991A-23216 Sequence 23216, A |
| 16 | 94.5 | 9.1 | 673 | 4 | US-09-252-991A-29219 Sequence 29219, A |
| 17 | 93.5 | 9.0 | 138 | 4 | US-09-252-991A-26931 Sequence 26931, A |
| 18 | 93.5 | 9.0 | 234 | 4 | US-09-252-991A-28985 Sequence 28985, A |
| 19 | 93.5 | 9.0 | 259 | 4 | US-09-252-991A-23311 Sequence 23311, A |
| 20 | 92 | 8.9 | 182 | 4 | US-09-252-991A-22314 Sequence 22314, A |
| 21 | 92 | 8.9 | 249 | 4 | US-09-252-991A-31926 Sequence 31926, A |
| 22 | 92 | 8.9 | 371 | 4 | US-09-252-991A-25006 Sequence 25006, A |
| 23 | 92 | 8.9 | 480 | 4 | US-09-252-991A-25416 Sequence 25416, A |
| 24 | 91.5 | 8.8 | 486 | 4 | US-09-252-991A-30369 Sequence 30369, A |
| 25 | 91.5 | 8.8 | 367 | 4 | US-09-252-991A-28247 Sequence 28247, A |
| 26 | 91 | 8.8 | 162 | 4 | US-09-252-991A-18207 Sequence 18207, A |
| 27 | 91 | 8.8 | 177 | 4 | US-09-252-991A-31950 Sequence 31950, A |

| | | | | | |
|----|------|-----|------|---|--|
| 28 | 91 | 8.8 | 279 | 4 | US-09-252-991A-26860 Sequence 26860, A |
| 29 | 91 | 8.8 | 433 | 4 | US-09-252-991A-28695 Sequence 28695, A |
| 30 | 91 | 8.8 | 447 | 4 | US-09-252-991A-27262 Sequence 27262, A |
| 31 | 90.5 | 8.7 | 220 | 4 | US-09-252-991A-16739 Sequence 16739, A |
| 32 | 90.5 | 8.7 | 253 | 4 | US-09-252-991A-29632 Sequence 29632, A |
| 33 | 90.5 | 8.7 | 631 | 3 | US-08-927-219-127 Sequence 127, App |
| 34 | 90.5 | 8.7 | 677 | 4 | US-09-252-991A-21541 Sequence 21541, A |
| 35 | 90.5 | 8.7 | 678 | 4 | US-09-252-991A-20406 Sequence 20406, A |
| 36 | 90.5 | 8.7 | 1487 | 2 | US-08-760-489-2 Sequence 2, Appl1 |
| 37 | 90.5 | 8.7 | 1487 | 2 | US-08-760-489-4 Sequence 4, Appl1 |
| 38 | 90.5 | 8.7 | 1487 | 3 | US-09-185-373-2 Sequence 2, Appl1 |
| 39 | 90.5 | 8.7 | 1487 | 3 | US-09-185-373-4 Sequence 4, Appl1 |
| 40 | 90 | 8.7 | 253 | 4 | US-09-252-991A-32459 Sequence 32459, A |
| 41 | 90 | 8.7 | 377 | 4 | US-09-252-991A-20479 Sequence 20479, A |
| 42 | 90 | 8.7 | 390 | 4 | US-09-252-991A-19933 Sequence 19933, A |
| 43 | 90 | 8.7 | 422 | 4 | US-09-252-991A-19452 Sequence 19452, A |
| 44 | 90 | 8.7 | 458 | 4 | US-09-252-991A-27645 Sequence 27645, A |
| 45 | 89.5 | 8.6 | 240 | 4 | US-09-252-991A-31100 Sequence 31100, A |

ALIGNMENTS

RESULT 1
US-09-878-281A-148
; Sequence 148, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 115
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-09-878-281A-148

Query/Match 31.0%; Score 322; DB 4; Length 115;
Best Local Similarity 62.0%; Pred. No. 3.4e-24;
Matches 67; Conservative 4; Mismatches 35; Indels 2; Gaps 1;

QY 7 KEKPNVTPTVAHRTSSRAVAVSLVEPTCCRAAGDWCARCRGRPSGRNLEVDVSLSPR 66
DB 10 KTKRNTNP--GRTLSQDAVAVSLVEFTCYHAGASWVCVOCARLPSGRNLAVGANPSFG 67
QY 67 HVPAPAGPGLSPGTLGPMARVAGARDGSCLPVALGLAGAPOTPGVG 114
DB 68 RAEPAPAGPGLSPGTLGPMARVAGARDGSCLPVALGLAGAPOTPGVG 115

RESULT 2
US-09-252-991A-27618
; Sequence 27618, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27618
; LENGTH: 236

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:52:49 ; Search time 207.892 Seconds
(without alignments)

542.461 Million cell updates/sec

```
Title: US-09-719-277C-2
Perfect score: 1039
Sequence: 1 ARILNKEKPNVTPTVAHRT.....WPCSLALCPLOPTKCAIPRG 196
```

Scoring table: BLOSUM62
GapPen 10.0 GapExt 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

| | |
|------------------|--------------------|
| Post-processing: | Minimum Match 0% |
| | Maximum Match 100% |

Listing first 45 summaries

```
Database : uniprot_02:*
1: uniprot_sprot:*
2: uniprot_tramb1:*
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match. | Length | DB | ID | Description |
|------------|-------|--------------|--------|----|----------|---------------------|
| 1 | 412 | 39.7 | 122 | 2 | 000687 | 000687 hepatitis c |
| 2 | 386 | 37.2 | 95 | 2 | 068874 | 068874 hepatitis c |
| 3 | 357.5 | 34.4 | 134 | 2 | 0700G2 | 0700d2 hepatitis c |
| 4 | 357.5 | 34.4 | 134 | 2 | CAR46584 | Ca66584 hepatitis c |
| 5 | 356 | 34.3 | 102 | 2 | 068358 | 068358 hepatitis c |
| 6 | 355 | 34.2 | 102 | 2 | 068351 | 068351 hepatitis c |
| 7 | 335.5 | 32.3 | 102 | 2 | 068356 | 068356 hepatitis c |
| 8 | 333 | 32.1 | 102 | 2 | 068359 | 068359 hepatitis c |
| 9 | 316 | 30.4 | 71 | 2 | 091AW2 | 091aw2 hepatitis c |
| 10 | 311 | 29.9 | 71 | 2 | 091AW1 | 091aw1 hepatitis c |
| 11 | 311 | 29.9 | 100 | 2 | 068360 | 068360 hepatitis c |
| 12 | 301 | 29.0 | 102 | 2 | 068365 | 068365 hepatitis c |
| 13 | 299 | 28.8 | 102 | 2 | 068366 | 068366 hepatitis c |
| 14 | 276 | 26.6 | 102 | 2 | 068363 | 068363 hepatitis c |
| 15 | 266 | 26.6 | 102 | 2 | 068368 | 068368 hepatitis c |
| 16 | 268 | 25.8 | 102 | 2 | 068367 | 068367 hepatitis c |
| 17 | 243 | 23.4 | 80 | 2 | 081289 | 081289 hepatitis c |
| 18 | 215.5 | 20.7 | 119 | 2 | 066686 | 066686 hepatitis c |
| 19 | 137 | 13.2 | 53 | 2 | 091KM9 | 091km9 hepatitis c |
| 20 | 114.5 | 11.0 | 210 | 2 | 06NY11 | 06ny11 homo sapiens |
| 21 | 114.5 | 11.0 | 210 | 2 | AAH65447 | Aah65547 homo sapi |
| 22 | 106.5 | 10.3 | 409 | 2 | 073V80 | 073v80 mycobacteri |
| 23 | 106.5 | 10.3 | 409 | 2 | AA505258 | Aa505258 mycobactr |
| 24 | 106 | 10.2 | 108 | 2 | P87760 | P87760 hepatitis c |
| 25 | 106 | 10.2 | 108 | 2 | P87761 | P87761 hepatitis c |
| 26 | 100 | 9.6 | 363 | 3 | 065949 | 065949 streptomyces |
| 27 | 100 | 9.6 | 926 | 2 | 08R554 | 08r554 mus musculu |
| 28 | 100 | 9.6 | 3651 | 2 | 08XK69 | 083x69 streptomyces |
| 29 | 99.5 | 9.5 | 394 | 2 | 097560 | 097560 macropus h |
| 30 | 99 | 9.5 | 356 | 2 | 09J314 | 09j314 aeromonas b |
| 31 | 98 | 9.4 | 443 | 2 | 09H718 | 09h718 homo sapiens |

| | | | | | | | |
|----|------|-----|------|---|------------|----------|--------------|
| 32 | 98 | 9.4 | 500 | 2 | Q6ZNL1 | Q6ZNL1 | homo sapien |
| 33 | 98 | 9.4 | 500 | 2 | BAC85133 | BAC85133 | homo sapi |
| 34 | 98 | 9.4 | 705 | 2 | Q82BP6 | Q82BP6 | streptomyces |
| 35 | 97.5 | 9.4 | 2205 | 2 | Q7PS10 | Q7PS10 | anophel |
| 36 | 97 | 9.3 | 382 | 2 | Q9U169 | Q9U169 | leishmania |
| 37 | 96.5 | 9.3 | 205 | 2 | Q7XK09 | Q7XK09 | oryza sativ |
| 38 | 96.5 | 9.3 | 205 | 2 | Q8W2P8 | Q8W2P8 | oryza sativ |
| 39 | 96.5 | 9.3 | 804 | 2 | Q6ZP11 | Q6ZP11 | mus muscul |
| 40 | 96.5 | 9.3 | 804 | 2 | BAC98254 | BAC98254 | mus muscul |
| 41 | 96 | 9.2 | 282 | 2 | Q7PRL8 | Q7PRL8 | anophel |
| 42 | 96 | 9.2 | 1191 | 2 | Q9RHV0 | Q9RHV0 | streptomyces |
| 43 | 95.5 | 9.2 | 250 | 2 | Q8N8C1 | Q8N8C1 | homo sapien |
| 44 | 95.5 | 9.2 | 514 | 1 | IMA_AATMEL | Q16992 | anthopleura |
| 45 | 95.5 | 9.2 | 802 | 2 | Q8K327 | Q8K327 | mus muscul |

RESULT 1

| ID | 000687; | PRELIMINARY; | PRT; | 122 AA. |
|----------|--|--------------|------|---------|
| AC | 01-NOV-1996 (TrEMBLrel. 01, Created) | | | |
| AD | 01-NOV-1996 (TrEMBLrel. 01, Last sequence update) | | | |
| DT | 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) | | | |
| DS | Core (Fragment). | | | |
| GN | Name=core; | | | |
| OS | Hepatitis C virus. | | | |
| OC | viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; | | | |
| OX | HepatitisB. | | | |
| NCBI | taxid:11103; | | | |
| UN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | MEDLINE=92268871; Pubmed=1316939; | | | |
| RA | Chan S., McCormick F., Holmes E., Dow B., Peutherer J., Follett E., | | | |
| RT | Yap P., Simmonds P. | | | |
| RL | "Analysis of a new hepatitis C virus type and its phylogenetic | | | |
| RL | relationship to existing variants." | | | |
| RL | J. Gen. Virol. 73:1131-1141(1992). | | | |
| UN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | MEDLINE=94270990; Pubmed=7545932; | | | |
| RA | Hochta H., Doi H., Hayashi T., Purwanta M., Soemarto W., Mizokami M. | | | |
| RA | Obba K., Homma M.; | | | |
| RT | "Analysis of the core and E1 envelope region sequences of a novel | | | |
| RT | variant of hepatitis C virus obtained in Indonesia." | | | |
| RL | Arch. Virol. 136:53-62(1994). | | | |
| UN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Chan S.-W.; | | | |
| RA | Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL, D10123; BAA01000.1; -. | | | |
| FT | NON_TER 1 1 | | | |
| FT | NON_TER 122 122 | | | |
| SEQUENCE | 122 AA; 12632 MW; FB5C7A1F76AE29AC CRC64; | | | |

Query Match

Best Local Similarity 69.4%; Pred.No. 1.8e-23;
Matches 84; Conservative 8; Mismatches 29; Indels 0; Gaps 0.

```
6 LKEXPNPTVARTSSSRVA VRSLSVEBTPCCRGALDWCARRRRLPSGNLEVDVSLSP 65
```

| | | | | |
|----|--|-----|---|-----|
| Oy | | 66 | RHVPRAGPGSLPPTLGPSSMMRWVAGSDGSCLPVALGLAIPOTPVGVRAIVRSSTPL | 125 |
| | | : | : : | : |
| | | : | : : | : |
| | | : | : : | : |
| | | : | : : | : |
| Dd | | 61 | KRVAKKGGPSLPSITRGSPWTRAAAGCGGSCPHAPVLALGAQMTVPGRGPAIWRVSSTIPS | 120 |
| | | : | : : | : |
| Oy | | 126 | R | 126 |
| | | : | : : | : |
| Dd | | 121 | R | 121 |

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 15:01:55 : Search time 2749 Seconds
(without alignments)
7887.101 Million cell updates/sec

Title: US-09-719-277C-1

Perfect score: 595
Sequence: 1 gacgaacatcctaactca.....gtgcgaatcctctg99gct 595

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----------|---------------------|
| 1 | 55.8 | 9.4 | 925 | CNS0091P | AL053013 Drosophila |
| 2 | 52.6 | 8.8 | 1201 | CNS0168R | AL106545 Drosophila |
| 3 | 47.6 | 8.0 | 839 | CNS0049B | AL054280 Drosophila |
| 4 | 46.6 | 7.8 | 925 | CNS0091P | AL053013 Drosophila |
| 5 | 46.4 | 7.8 | 1010 | CNS0167E | AL107180 Drosophila |
| 6 | 45.4 | 7.6 | 844 | CNS0052P | AL056652 Drosophila |
| 7 | 43.8 | 7.4 | 910 | CNS0060N | AL056652 Drosophila |
| 8 | 43.8 | 7.4 | 1100 | CNS0168D | AL106855 Drosophila |
| 9 | 43.4 | 7.3 | 1101 | CNS0175Y | AL108460 Drosophila |
| 10 | 43.2 | 7.3 | 411 | CC051731 | CC051731 SALI_0111 |
| 11 | 43.2 | 7.3 | 645 | CNS01213 | AL101589 Drosophila |
| 12 | 43.2 | 7.3 | 1009 | CNS0108W | AL098882 Drosophila |
| 13 | 42.8 | 7.2 | 692 | CNS0070H | AL050923 Drosophila |
| 14 | 42.6 | 7.2 | 775 | AZ184156 | AZ184156 SP_1002.B |
| 15 | 42.6 | 7.2 | 912 | CNS006N3 | AL065775 Drosophila |
| 16 | 42.6 | 7.1 | 1009 | CNS0108W | AL098882 Drosophila |
| 17 | 41.2 | 6.9 | 630 | AQ257510 | AQ257510 nxb00181 |
| 18 | 41.2 | 6.9 | 977 | CNS00JX7 | AL076850 Drosophila |
| 19 | 41.2 | 6.9 | 1190 | B0673485 | B0673485 AGENCOURT |
| 20 | 41.1 | 6.9 | 356 | BY235468 | BY235468 Tetrarodon |
| 21 | 41.1 | 6.9 | 879 | CNS0200G | AL209545 Tetrarodon |
| 22 | 41.1 | 6.9 | 936 | BG852371 | BG852371 Drosophila |
| 23 | 41.1 | 6.9 | 1101 | CNS008NG | AL057398 Drosophila |
| 24 | 40.8 | 6.9 | 1036 | CNS0108S | AL098770 Drosophila |

| | | | | | | |
|------|------|-----|------|---|----------|---------------------|
| C 25 | 40.6 | 6.8 | 754 | 9 | AG054378 | AG054378 Pan trogl |
| C 26 | 40.6 | 6.8 | 893 | 9 | CL468979 | CL468979 SAHL_1295 |
| C 27 | 40.6 | 6.8 | 967 | 9 | CG294358 | CG294358 OG2B76TV |
| C 28 | 40.6 | 6.8 | 995 | 9 | CG244380 | CG244380 OG1A82TV |
| C 29 | 40.6 | 6.8 | 1007 | 8 | BZ569268 | BZ569268 pac82-164 |
| C 30 | 40.6 | 6.8 | 970 | 9 | CNS010C9 | AL098787 Drosophila |
| C 31 | 40.2 | 6.8 | 932 | 9 | CNS00720 | AL066742 Drosophila |
| C 32 | 40.2 | 6.8 | 1101 | 9 | CNS0167E | AL107216 Drosophila |
| C 33 | 40.2 | 6.7 | 846 | 9 | CNS0108J | AL099337 Drosophila |
| C 34 | 39.8 | 6.7 | 942 | 1 | AF367693 | AF367693 AF367693 |
| C 35 | 39.8 | 6.7 | 972 | 5 | BQ900341 | BQ900341 AGENCOURT |
| C 36 | 39.8 | 6.7 | 990 | 5 | BUI63350 | BUI63350 AGENCOURT |
| C 37 | 39.8 | 6.7 | 1135 | 5 | BUS27635 | BUS27635 AGENCOURT |
| C 38 | 39.6 | 6.7 | 664 | 8 | BH977944 | BH977944 cde92e01 |
| C 39 | 39.6 | 6.7 | 1201 | 9 | CNS015M3 | AL105981 Drosophila |
| C 40 | 39.6 | 6.7 | 1475 | 9 | CL469730 | CL469730 SAHL_133 |
| C 41 | 39.4 | 6.6 | 411 | 9 | CG569739 | CG569739 OST18057 |
| C 42 | 39.4 | 6.6 | 772 | 8 | AO740673 | AO740673 HS_5507 |
| C 43 | 39.4 | 6.6 | 798 | 5 | BP144487 | BP144487 BRL44487 |
| C 44 | 39.4 | 6.6 | 931 | 2 | BP122031 | BP122031 601756375 |
| C 45 | 39.4 | 6.6 | 955 | 9 | CNS008SF | AL069794 Drosophila |

ALIGNMENTS

RESULT 1
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BAC19016 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL053013
VERSION AL053013.1 GI:4934461
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE Drosophila melanogaster
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
JOURNAL 1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC19016"
/clone_lib="RPI-98"
/note="end : TET3"

ORIGIN

Query Match 9.4%; Score 55.8; DB 9; Length 925;
Best Local Similarity 14.2%; Pred. No. 0.0066;

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 15:06:26 ; Search time 88 Seconds

(without alignments)
4805.902 Million cell updates/sec

Title: US-09-719-277C-1

Perfect score: 595
Sequence: 1 gcagcaatcccaactcaaa.....gtgcgaattctctcgggact 595

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 segs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--|
| 1 | 595 | 100.0 | 9416 | 4 | US-08-823-895A-26 Sequence 26, Appl |
| 2 | 595 | 100.0 | 9416 | 4 | US-10-104-966-13 Sequence 13, Appl |
| 3 | 595 | 100.0 | 9539 | 3 | US-09-014-416-2 Sequence 2, Appl |
| 4 | 595 | 100.0 | 9646 | 3 | US-08-811-566-1 Sequence 1, Appl |
| 5 | 595 | 100.0 | 9646 | 3 | US-09-034-756-1 Sequence 1, Appl |
| 6 | 593.4 | 99.7 | 12980 | 3 | US-08-811-566-5 Sequence 5, Appl |
| 7 | 593.4 | 99.7 | 12980 | 3 | US-09-034-756-5 Sequence 5, Appl |
| 8 | 591.8 | 99.5 | 9401 | 2 | US-08-432-693-1 Sequence 1, Appl |
| 9 | 591.8 | 99.5 | 9416 | 3 | US-08-811-566-19 Sequence 19, Appl |
| 10 | 591.8 | 99.5 | 9416 | 3 | US-09-034-756-19 Sequence 19, Appl |
| 11 | 582.6 | 97.9 | 9401 | 5 | PCT-US91-02225-9 Sequence 9, Appl |
| 12 | 581 | 97.9 | 9401 | 5 | US-08-444-818-176 Sequence 176, App |
| 13 | 580.6 | 97.6 | 657 | 2 | US-07-681-703B-11 Sequence 11, Appl |
| 14 | 580.6 | 97.6 | 657 | 2 | US-08-407-410B-11 Sequence 11, Appl |
| 15 | 580.6 | 97.6 | 657 | 2 | US-08-485-500-11 Sequence 11, Appl |
| 16 | 580.6 | 97.6 | 657 | 5 | PCT-US91-02370-11 Sequence 11, Appl |
| 17 | 580.6 | 97.6 | 9185 | 3 | US-08-444-818-122 Sequence 122, App |
| 18 | 580.6 | 97.6 | 9185 | 3 | US-08-444-818-123 Sequence 123, App |
| 19 | 580.6 | 97.6 | 9379 | 4 | US-09-388-874-1 Sequence 1, Appl |
| 20 | 580.6 | 97.6 | 9379 | 4 | US-09-916-359-1 Sequence 1, Appl |
| 21 | 580.6 | 97.6 | 9401 | 1 | US-07-910-760-9 Sequence 9, Appl |
| 22 | 580.6 | 97.6 | 9401 | 1 | US-08-440-519-9 Sequence 9, Appl |
| 23 | 580.6 | 97.6 | 9401 | 4 | US-08-440-549-9 Sequence 9, Appl |
| 24 | 580.6 | 97.6 | 9401 | 4 | US-08-823-895A-25 Sequence 25, Appl |
| 25 | 579 | 97.3 | 8987 | 1 | US-08-444-818-137 Sequence 137, App |
| 26 | 575.8 | 96.8 | 657 | 1 | US-07-681-703B-13 Sequence 13, Appl |
| 27 | 575.8 | 96.8 | 657 | 2 | US-08-407-410B-13 Sequence 13, Appl |

| | | | | | |
|----|-------|------|------|---|---|
| 28 | 575.8 | 96.8 | 657 | 2 | US-08-485-500-13 Sequence 13, Appl |
| 29 | 575.8 | 96.8 | 657 | 2 | PCT-US91-02370-13 Sequence 16, Appl |
| 30 | 575.8 | 96.8 | 1539 | 2 | US-08-470-426B-16 Sequence 13, Appl |
| 31 | 575.8 | 96.8 | 1863 | 2 | US-08-470-426B-13 Sequence 104, App |
| 32 | 564.2 | 94.8 | 573 | 5 | US-08-290-665A-104 Sequence 104, App |
| 33 | 564.2 | 94.8 | 573 | 5 | PCT-US95-10398-104 Sequence 1, Appl |
| 34 | 559 | 93.9 | 630 | 3 | US-09-203-649-1 Sequence 1, Appl |
| 35 | 559 | 93.9 | 630 | 4 | US-09-968-253-1 Sequence 103, App |
| 36 | 557.8 | 93.7 | 573 | 2 | US-08-290-665A-103 Sequence 103, App |
| 37 | 557.8 | 93.7 | 573 | 2 | PCT-US95-10398-103 Sequence 105, App |
| 38 | 554.6 | 93.2 | 573 | 5 | US-08-290-665A-105 Sequence 105, App |
| 39 | 554.6 | 93.2 | 573 | 5 | PCT-US95-10398-105 Sequence 106, App |
| 40 | 551.4 | 92.7 | 573 | 2 | US-08-290-665A-106 Sequence 106, App |
| 41 | 551.4 | 92.7 | 573 | 5 | PCT-US95-10398-106 Sequence 107, App |
| 42 | 544 | 91.4 | 573 | 2 | US-08-290-665A-107 Sequence 107, App |
| 43 | 544 | 91.4 | 573 | 5 | PCT-US95-10398-107 Sequence 108, App |
| 44 | 541.8 | 91.1 | 573 | 2 | US-08-290-665A-108 Sequence 108, App |
| 45 | 541.8 | 91.1 | 573 | 5 | PCT-US95-10398-108 Sequence 108, App |

ALIGNMENTS

RESULT 1
US-08-823-895A-26
Sequence 26, Application US/08823895A
Patent No. 6433159
GENERAL INFORMATION:
APPLICANT: Kevin P. Anderson
TITLE OF INVENTION: Compositions And Methods For
Treatment Of Hepatitis C Virus-Associated Diseases
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Jane Massey Licata, Esq.
STREET: 66 E. Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823, 895A
FILING DATE: March 17, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/453, 085
FILING DATE: May 30, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/945, 289
FILING DATE: September 10, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416
TYPE: Nucleic
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
US-08-823-895A-26

Query Match 100.0%; Score 595; DB 4; Length 9416;
Best Local Similarity 100.0%; Pred. No. 2.8e-161;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

/notes="Xaa is unspecified and corresponds to in-frame
stop codon"
99WO-US012929.
98US-0088670P.
98US-0089138P.
A D.
CT J L.
) D.
ewaki JL, Stump DD,
1/11.
5.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 13:30:50 ; Search time 3013 Seconds
(without alignments)
9338.656 Million cell updates/sec

Title: US-09-719-277C-1

Perfect score: 595
Sequence: 1 gcaacgaatcctaaccctca.....gtcgcaatcctcggggc 595

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 2364489745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

GenBank:
1: gb_ba:
2: gb_ncg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sbs:
12: gb_ey:
13: gb_un:
14: gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|------------------|--------------------|
| 1 | 595 | 100.0 | 2610 14 HPCST77 | M62381 Hepatitis C |
| 2 | 595 | 100.0 | 9416 6 AR453138 | AR453138 Sequence |
| 3 | 595 | 100.0 | 9416 6 AX441173 | AX441173 Sequence |
| 4 | 595 | 100.0 | 9416 14 HPCCGA | M67463 Hepatitis C |
| 5 | 595 | 100.0 | 9518 6 AX100563 | AX100563 Sequence |
| 6 | 595 | 100.0 | 9599 6 AR119831 | AR119831 Sequence |
| 7 | 595 | 100.0 | 9599 14 AF011751 | AF011751 Hepatitis |
| 8 | 595 | 100.0 | 9599 14 AF011752 | AF011752 Hepatitis |
| 9 | 595 | 100.0 | 9599 14 AF011753 | AF011753 Hepatitis |
| 10 | 595 | 100.0 | 9646 6 AR110828 | AR110828 Sequence |
| 11 | 595 | 100.0 | 9646 6 BD069882 | BD069882 Functiona |
| 12 | 595 | 100.0 | 9646 14 AF009606 | AF009606 Hepatitis |
| 13 | 595 | 100.0 | 12119 6 AX057315 | AX057315 Sequence |
| 14 | 593.4 | 99.7 | 12980 6 AR110831 | AR110831 Sequence |
| 15 | 593.4 | 99.7 | 12980 6 BD069885 | BD069885 Functiona |
| 16 | 591.8 | 99.5 | 978 14 HPCCEXPA | M55970 Human hepat |
| 17 | 591.8 | 99.5 | 9365 14 AF290978 | AF290978 Hepatitis |
| 18 | 591.8 | 99.5 | 9401 6 AR030378 | AR030378 Sequence |
| 19 | 591.8 | 99.5 | 9416 6 AR110845 | AR110845 Sequence |

| | | | | | |
|----|-------|------|---------|-----------|--------------------|
| 20 | 591.8 | 99.5 | 9416.6 | BD069999 | BD069999 Functiona |
| 21 | 591.8 | 99.5 | 9622.6 | AX663428 | AX663428 Sequence |
| 22 | 587 | 98.7 | 9618.14 | AF271632 | AF271632 Hepatitis |
| 23 | 585.4 | 98.4 | 1805.14 | AY231582 | AY231582 Hepatitis |
| 24 | 583.8 | 98.1 | 1288.14 | AF268569 | AF268569 Hepatitis |
| 25 | 583.8 | 98.1 | 1288.14 | AF268570 | AF268570 Hepatitis |
| 26 | 583.8 | 98.1 | 1288.14 | AF268571 | AF268571 Hepatitis |
| 27 | 583.8 | 98.1 | 1620.14 | HPCNSISPH | M74811 Hepatitis C |
| 28 | 583.8 | 98.1 | 1626.14 | HPCNSISPE | M74808 Hepatitis C |
| 29 | 583.8 | 98.1 | 2610.14 | HPCST90 | M62382 Hepatitis C |
| 30 | 583.8 | 98.1 | 9609.12 | AF387805 | AF387805 Synthetic |
| 31 | 583.8 | 98.1 | 9609.12 | AF387806 | AF387806 Synthetic |
| 32 | 583.8 | 98.1 | 9646.12 | AF387807 | AF387807 Synthetic |
| 33 | 583.8 | 98.1 | 9693.12 | AF387807 | AF387807 Synthetic |
| 34 | 582.2 | 97.8 | 1288.14 | AF268572 | AF268572 Hepatitis |
| 35 | 582.2 | 97.8 | 1626.14 | HPCNSISPI | M74812 Hepatitis C |
| 36 | 581 | 97.6 | 9379.6 | AR118747 | AR118747 Sequence |
| 37 | 580.6 | 97.6 | 657.6 | AR061075 | AR061075 Sequence |
| 38 | 580.6 | 97.6 | 657.6 | AR061088 | AR061088 Sequence |
| 39 | 580.6 | 97.6 | 657.6 | 113940 | 113940 Sequence 11 |
| 40 | 580.6 | 97.6 | 657.6 | AX032707 | AX032707 Sequence |
| 41 | 580.6 | 97.6 | 657.6 | BD078641 | BD078641 Hepatitis |
| 42 | 580.6 | 97.6 | 9185.6 | AR118722 | AR118722 Sequence |
| 43 | 580.6 | 97.6 | 9185.6 | AR118723 | AR118723 Sequence |
| 44 | 580.6 | 97.6 | 9185.6 | 108294 | 108294 Sequence 1 |
| 45 | 580.6 | 97.6 | 9185.6 | BD091382 | BD091382 HCV culti |

ALIGNMENTS

| | | | | |
|------------|--|--|--------|-----------------|
| RESULT 1 | HPCST77 | 2610 bp ss-RNA | linear | VRL 16-OCT-2002 |
| LOCUS | HPCST77 | Hepatitis C virus polyprotein gene, partial cds. | | |
| DEFINITION | Hepatitis C virus polyprotein gene, partial cds. | | | |
| ACCESSION | M62381 | | | |
| VERSION | M62381.1 GI:329972 | | | |
| KEYWORDS | structural protein. | | | |
| SOURCE | Hepatitis C virus | | | |
| ORGANISM | Hepatitis C virus | | | |
| REFERENCE | 1 (bases 1 to 2610) | | | |
| AUTHORS | Ogata,N., Alter,H.J., Miller,R.H. and Purcell,R.H. | | | |
| TITLE | Nucleotide sequence and mutation rate of the H strain of hepatitis C virus | | | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 88 (8), 3392-3396 (1991) | | | |
| MEDLINE | 91195357 | | | |
| PubMed | 1849654 | | | |
| FEATURES | location/Qualifiers | | | |
| source | 1..2610 | | | |
| | /organism="Hepatitis C virus" | | | |
| | /mol_type="genomic RNA" | | | |
| | /specific_host="Patient 'H' 1977" | | | |
| | /db_xref="taxon:11103" | | | |
| | 277..>2610 | | | |
| | /note="largest ORF" | | | |
| | /product="polyprotein" | | | |
| | /protein_id="AA02127.1" | | | |
| | /translation="MSTNPKRQKTKNTNRRODVFPGGQIVGVVLLPRGRPL GVATRTSRSGRGRORIPKARREGRTWAOGPVWVAGAGAGGMLSPG SRPWSGPTDPRRSRNIGKIDITLGGFADIMGIVLVAIGAGAGGVVLD GVNATGNGPCSFSTLALISGLTPPAAYVVRNSGIVHTNDPNSIYRBA AIIHTPECPVCRGNARCVAVTPVATRDGLPTQLRHIDLVGATCSALV VBDLGSVPLVQLFTSPRRHTTDCNSIYGFVSWGMAKVLVILLFAGDAETHVGS AQLRIPOALMDIAGAMVLAGIAFVSVMGAKVLVILLFAGDAETHVGS GRTTAGVGLTPOAKONILQVNSGMLNSTANGSLINGLAGLYRHHKNS GCEPLRACGRITDPAQGMGIVTANGSGIDRKYCMHYVPRGQIVPAASVGGPV FTSPVYVGTDSGAPITYSWANDVDVPLANTRPPLGWRFGCTKMSGPTKVC PCVLTGGVNTLLCPDCKRKPATYSKSGSPWITPSCMDVDFRLMHPCTLY TIRKRYVGVGVEHRLAECNMTGRCDLDRDRSLTLLSTTQOVLPCSFYTL | | | |